

University of Groningen

## Cosmology with intensity mapping techniques using atomic and molecular lines

Fonseca, José; Silva, Marta B.; Santos, Mário G.; Cooray, Asantha

*Published in:*  
Monthly Notices of the Royal Astronomical Society

*DOI:*  
[10.1093/mnras/stw2470](https://doi.org/10.1093/mnras/stw2470)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. (2017). Cosmology with intensity mapping techniques using atomic and molecular lines. *Monthly Notices of the Royal Astronomical Society*, 464(2), 1948-1965. <https://doi.org/10.1093/mnras/stw2470>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Cosmology with intensity mapping techniques using atomic and molecular lines

José Fonseca,<sup>1★</sup> Marta B. Silva,<sup>2</sup> Mário G. Santos<sup>1,3,4</sup> and Asantha Cooray<sup>5</sup>

<sup>1</sup>Department of Physics and Astronomy, University of the Western Cape, Cape Town 7535, South Africa

<sup>2</sup>Kapteyn Astronomical Institute, University of Groningen, Landleven 12, NL-9747 AD Groningen, the Netherlands

<sup>3</sup>SKA SA, The Park, Park Road, Cape Town 7405, South Africa

<sup>4</sup>CENTRA, Instituto Superior Técnico, Universidade de Lisboa, P-1049-001 Lisboa, Portugal

<sup>5</sup>Department of Physics & Astronomy, University of California, Irvine, CA 92697, USA

Accepted 2016 September 27. Received 2016 September 23; in original form 2016 July 20

## ABSTRACT

We present a systematic study of the intensity mapping (IM) technique using updated models for the different emission lines from galaxies. We identify which ones are more promising for cosmological studies of the post-reionization epoch. We consider the emission of Ly $\alpha$ , H $\alpha$ , H $\beta$ , optical and infrared oxygen lines, nitrogen lines, C II and the CO rotational lines. We show that Ly $\alpha$ , H $\alpha$ , O II, C II and the lowest rotational CO lines are the best candidates to be used as IM probes. These lines form a complementary set of probes of the galaxies' emission spectra. We then use reasonable experimental setups from current, planned or proposed experiments to assess the detectability of the power spectrum of each emission line. IM of Ly $\alpha$  emission from  $z = 2$  to 3 will be possible in the near future with Hobby–Eberly Telescope Dark Energy Experiment, while far-infrared lines require new dedicated experiments. We also show that the proposed *SPHEREx* satellite can use O II and H $\alpha$  IM to study the large-scale distribution of matter in intermediate redshifts of 1–4. We find that submillimetre experiments with bolometers can have similar performances at intermediate redshifts using C II and CO(3–2).

**Key words:** cosmology; miscellaneous – large-scale structure of Universe.

## 1 INTRODUCTION

Measurements of the three-dimensional (3D) large-scale structure of the Universe across cosmic time promise to bring exquisite constraints on cosmology, from the nature of dark energy or the mass of neutrinos to primordial non-Gaussianity and tests of General Relativity (GR) on large scales. Most of these surveys are based on imaging a large number of galaxies at optical or near-infrared (NIR) wavelengths, combined with redshift information to provide 3D positions of the galaxies [BOSS (SDSS-III) (Schlegel, White & Eisenstein 2009), DES (Flaugher 2005), eBOSS (Dawson et al. 2016), DESI (Levi et al. 2013), 4MOST (de Jong et al. 2012), LSST (LSST Dark Energy Science Collaboration 2012; Bacon et al. 2015), WFIRST (Spergel et al. 2015) and the *Euclid* satellite (Laureijs et al. 2011; Amendola et al. 2013)]. These observations at optical and infrared (IR) wavelengths will be limited to galaxy samples between  $z = 0.3$  and 2, and in some cases, redshifts will be determined only by the photometric data.

Instead of counting galaxies, the intensity mapping (IM) technique uses the total observed intensity from any given pixel.

For a reasonably large 3D pixel (with a given angular and frequency/redshift resolution), also referred to as *voxel*, we expect it to contain several galaxies. The intensity in each pixel will thus be the integrated emission from all these galaxies. This should then provide a higher signal-to-noise (SNR) compared to standard galaxy ‘threshold’ surveys. Moreover, since most cosmological applications rely on probing large scales, the use of these large pixels will not affect the cosmological constraints. By not needing galaxy detections, the requirements on the telescope/survey will be much less demanding. However, since we are no longer relying on ‘clean’ galaxy counts, we need to be much more careful with other contaminants of the observed intensity.

In order to have redshift information, the measured intensity should originate from specific emission lines. The underlying idea is that the amplitude of this intensity will be related to the number of galaxies in the 3D pixel emitting the target line. The fluctuations in the intensity across the map should then be proportional to the underlying dark matter (DM) fluctuations. Several lines can in principle be used for such surveys. In particular, a significant focus has been given in recent years to the H I 21 cm line (Battye, Davies & Weller 2004; Chang et al. 2008; Loeb & Wyithe 2008; Bagla, Khandai & Datta 2010; Ansari et al. 2012; Switzer et al. 2013) and how well it can perform cosmological measurements (see e.g. Camera et al.

\* E-mail: josecarlos.s.fonseca@gmail.com

2013; Bull et al. 2015). The 21 cm signal is also being used to study the epoch of reionization (EoR) at higher redshifts, with several experiments running or in development (see e.g. Mellema et al. 2013 for a review). As this line will be observed at radio frequencies, telescopes will naturally have lower resolutions, making this line an obvious application for IM. These telescopes also usually have large fields of view (FoV) that allows them to cover large areas of the sky more quickly. Moreover, given the low frequencies, the H I line has very little background and foreground contamination from other lines, which is an advantage when comparing to IM with other lines as we will discuss later. Still, all lines have significant Galactic foregrounds that need to be removed or avoided. Several experiments have been proposed so far (Chen 2012; CHIME Collaboration 2012; Battye et al. 2013; Bigot-Sazy et al. 2016; Newburgh et al. 2016), and some observations have already been made (Kerp et al. 2011), including a detection in cross-correlation (Chang et al. 2010). Moreover, a large H I IM survey has been proposed for SKA1-MID (Santos et al. 2015).

Although IM with the H I line holds great promises for cosmology, it is still relevant to ask if we can use other lines for such purposes. An obvious reason is that the 21 cm line is quite weak. Although galaxy surveys with other lines have become routine, with detections made up to very high redshifts, the highest redshift detection for H I is  $z = 0.376$  (Fernández et al. 2016). Some studies have already been conducted for other lines, in particular for the EoR, such as Ly $\alpha$  (Silva et al. 2013), C II (Gong et al. 2012), CO (Gong et al. 2011a), H<sub>2</sub> (Gong, Cooray & Santos 2013) and others (Visbal & Loeb 2010). At lower redshifts, Pullen, Dore & Bock (2014) have considered Ly $\alpha$  to probe the underlying DM power spectrum in a fiducial experiment, although they assume a very strong signal coming from the intergalactic medium (IGM). In fact, Ly $\alpha$  IM has already been detected in cross-correlation with quasars (Croft et al. 2016). Uzgil et al. (2014) considered the 3D power spectrum of C II and other far-infrared (FIR) emission lines. Similarly, Pullen et al. (2013) used the large-scale structure distribution of matter to study CO emission while Breyse, Kovetz & Kamionkowski (2014) studied the possibility of using the first rotational transition of CO to do a survey at  $z \sim 3$ .

This paper presents a systematic study of all lines (besides H I) that can in principle be used for IM with reasonable experimental setups. It also compares them to determine which are the optimal lines to target for IM. It uses the latest observational and simulation results to predict the strength of the intensity and the expected bias towards DM. It then discusses the feasibility of surveys with these lines and which ones are more appropriate for cosmological applications. This is particularly relevant as the combination of different lines can bring exquisite constraints on large-scale effects, such as primordial non-Gaussianity, using the multi-tracer technique (Seljak 2009; Alonso & Ferreira 2015; Fonseca et al. 2015). Moreover, the cross-correlation can be particularly useful in dealing with foregrounds/backgrounds as well as systematics. Finally, even if such IM surveys are not extremely competitive for cosmology, the simple detection of such a signal and its power spectrum will bring invaluable information about the astrophysical processes involved in the production of such lines and the clustering of the corresponding galaxies.

We start with a review of the basic calculations for the average observed intensity of an emission line in Section 2. In Section 3, we model line emission in terms of the star formation rate (SFR) of a given halo and estimate the signal for different lines. We then compare lines in Section 4 so that we have an indication of which surveys to target in Section 5. The two following sections are

focused on briefly assessing other sources of emission/contamination and address how to deal with such issues. We conclude in Section 8.

## 2 LINE IM – REVIEW

### 2.1 Coordinates and volume factors

Let us consider a volume in space with comoving centre  $\chi_0$ , two angular directions perpendicular to the line of sight, and another direction along the line of sight. Using the flat sky approximation, the comoving coordinates of a point inside such volume will be

$$\mathbf{r} = \chi_0 + \Delta\theta_1 D_A \hat{\theta}_1 + \Delta\theta_2 D_A \hat{\theta}_2 + \Delta\nu \tilde{\gamma} \hat{r}. \quad (1)$$

The first two components are the displacements perpendicular to the line of sight, where  $D_A$  is the angular diameter distance in comoving units (for a flat universe it is just the comoving distance  $\chi$ ). The last component corresponds to the line of sight or the redshift/wavelength/frequency direction. A small variation in the comoving distance corresponds to a variation in the observed frequency as  $\Delta\chi = \tilde{\gamma} \Delta\nu$ , where  $\tilde{\gamma}$  is defined by

$$\tilde{\gamma} \equiv \frac{d\chi}{d\nu} = \frac{\lambda_e (1+z)^2}{H(z)}. \quad (2)$$

Here  $\lambda_e$  is the wavelength of the emitted photon and  $H$  is the Hubble parameter at the redshift  $z$ , so that  $\lambda_0 = \lambda_e(1+z)$ .

### 2.2 Average observed signal

Let us assume that galactic emission lines have a sharp line profile  $\psi(\lambda)$  that can be approximated by a Dirac delta function around a fixed wavelength  $\lambda$ , i.e.  $\psi(\lambda) \simeq \delta^D(\lambda_e - \lambda)$ . This way we can assume a one-to-one relationship between the observed wavelength and redshift. This approximation is valid as long as the observational pixel is considerably larger than the full width at half-maximum (FWHM) of the targeted line. Although this will generally be the case in IM experiments, we will discuss the implications of a non-sharp line profile in Section 2.3.

The average observed intensity of a given emission line is given by

$$\bar{I}_\nu(z) = \int_{M_{\min}}^{M_{\max}} dM \frac{dn}{dM} \frac{L(M, z)}{4\pi D_L^2} \tilde{\gamma} D_A^2, \quad (3)$$

where  $dn/dM$  (Haloes/Mpc<sup>3</sup>/M<sub>⊙</sub>) is the comoving halo mass function (HMF; Sheth & Tormen 1999),  $D_L$  is the luminosity distance ( $D_L = (1+z)\chi$  for a flat universe) and  $\tilde{\gamma} D_A^2$  is the comoving unit volume of the voxel. Note that this is the physical average intensity, and with the definition of  $\tilde{\gamma}$  (equation 2), its units are power per unit area per unit solid angle per unit frequency. The galaxy line luminosity  $L(M, z)$  depends on the halo mass and often evolves with redshift. The relevant mass interval of integration depends on the chosen line and is redshift dependent. We assume  $10^8 M_\odot$  as the minimum mass of a halo capable of having stellar formation due to atomic cooling. If we instead assumed stellar formation through molecular cooling, then the minimum mass would be considerably lower (down to  $\sim 10^6 M_\odot$ ; Visbal et al. 2014). If not otherwise stated, we will consider the DM haloes' mass in the range  $[10^8, 10^{15}] M_\odot$ .

Alternatively, one can estimate the average intensity of a line using luminosity functions (LF) via

$$\bar{I}_\nu(z) = \int_{L_{\min}/L^*}^{L_{\max}/L^*} d\left(\frac{L}{L^*}\right) \frac{L}{4\pi D_L^2} \phi(L) \tilde{\gamma} D_A^2. \quad (4)$$

We take the LF  $\phi$  to have the Schechter form

$$\phi(L) d\left(\frac{L}{L^*}\right) = \phi^* \left(\frac{L}{L^*}\right)^\alpha e^{-L/L^*} d\left(\frac{L}{L^*}\right), \quad (5)$$

where  $\alpha$  is the slope of the faint end,  $L^*$  is the turnover luminosity for the bright end decay and  $\phi^*$  is the overall normalization of the LF. The set of parameters  $\{\phi^*, L^*, \alpha\}$  depend on the line in consideration as well as the integration limits.

### 2.3 Expected signal from broad emission lines

In some cases, the considered emission line has a broad line profile due to the motion of the gas in the galaxy. This causes the previous discussion to break down as the signal from a galaxy will be observed at several pixels along a line of sight. To minimize this issue, one could consider frequency bins that are at least of the same order of magnitude as the FWHM of the line. Alternatively, one could estimate the amount of signal observed at each pixel using the line profile. In this case, the galaxy will be responsible for the flux observed in several pixels. Let us assume that we have a galaxy at redshift  $z_c$  and that the line in study has total luminosity  $L$  and a line profile  $\psi(\lambda_c)$ . Since  $\int \psi(\lambda_c) d\lambda_c = 1$ , we can write the luminosity dependence as

$$L_{\text{line}}(\lambda_c) = L \psi(\lambda_c). \quad (6)$$

Due to the line profile, galaxies at different redshifts will contribute to the observed signal at a given wavelength. Therefore, the intensity at an observed wavelength will be an integral over all redshifts weighted by the line profile. Hence, the expected observed flux will be

$$\bar{I}_v(\lambda_O) = \int dz \frac{\lambda_O}{(1+z)^2} \int_{M_{\min}}^{M_{\max}} dM \frac{dn}{dM} \times \frac{L(M, z) \psi(\lambda_O/(1+z))}{4\pi D_L^2(z)} \bar{y}(z) D_A^2(z). \quad (7)$$

Such broadening of the emission line washes away small-scale fluctuations but one would not expect large effects on large angular scales. For the purpose of this paper, we will consider redshift bins larger than the broadening of the line, effectively approximating the emission line profile by a delta function.

## 3 ATOMIC AND MOLECULAR EMISSION LINES FROM GALAXIES

### 3.1 Relation between line luminosity and SFR

Our aim is to study the power spectrum of matter perturbations up to redshift  $z = 5$  by performing IM of the strongest emission lines from galaxies. Most of the lines we consider are sourced by ultraviolet (UV) emission from young stars and quasars. A fraction of the UV emission is absorbed by galactic dust that in turn will produce a strong continuum thermal emission that contributes to the IR continuum background. Another fraction of the ionizing UV radiation produced by galaxies escapes into the IGM, while a further part of this radiation is absorbed by the galaxy gas, heating it and exciting and/or ionizing its atoms and molecules. As the gas recombines and/or de-excites, transition lines are emitted from the gas. In most cases, the recombination time is significantly shorter than a Hubble time so we can assume that the process is immediate. As a result, the emission rate of a recombining line will be proportional to the stellar ionizing flux, the source of which is the star formation in

the galaxy. It therefore relates to the SFR, i.e. the amount of mass transformed into stars per year. This dependence is not necessarily linear and needs to be calibrated for each line. We further need to consider that only a portion of the line luminosity exits the galaxy. We then model the total luminosity associated with a halo of mass  $M$  as

$$L = K(z) \times \left( \frac{\text{SFR}(M, z)}{M_\odot \text{ yr}^{-1}} \right)^\gamma, \quad (8)$$

where  $K(z)$  includes the calibration of the lines as well as the several escape fractions, while the parameter  $\gamma$  is introduced for those lines that do not scale linearly with the SFR. This generic model was chosen to accommodate all the details of the lines that we will study in this paper.

The star formation rate density (SFRD) is defined as

$$\text{SFRD}(z) \equiv \int dM \text{SFR}(M, z) \frac{dn}{dM}, \quad (9)$$

where the integration is performed over all DM haloes and has units of  $M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$ . In this study, we attempt to cover the uncertainty in the SFRD by using two different SFR models.

The first model we consider uses the Markov chain Monte Carlo method to estimate the galaxies' stellar mass versus galaxy mass relation as a function of redshift based on several observational constraints in the literature (Behroozi, Wechsler & Conroy 2013). We will refer to this model as Be13. They provide the SFR as a function of halo mass and redshift, which can be found and downloaded from the authors' personal page.<sup>1</sup> They also studied the literature to produce a new best fit to the SFRD which reads

$$\text{SFRD}(z) = \frac{0.180}{10^{-0.997(z-1.243)} + 10^{0.241(z-1.243)}}. \quad (10)$$

Please note that when we computed the SFRD with Be13  $\text{SFR}(M, z)$  and the HMF using equation (9), we found a discrepancy between Be13 best fit to the SFRD and our estimate to the SFRD using Be13  $\text{SFR}(M, z)$ . Hence, we corrected the SFR by a redshift-dependent fraction such that we recover equation (10) when using our HMF in equation (9). We use a second model to account for the large uncertainty in SFR observations at high redshifts ( $z > 2$ ). This model is a parametrization of simulated galaxy catalogues, which we will refer to as SMill. The galaxy catalogues used were obtained by De Lucia & Blaizot (2007) and Guo et al. (2011) who post-processed the Millennium I (Springel et al. 2005) and II (Boylan-Kolchin et al. 2009) simulations, respectively. The SFR parametrization in terms of the mass and redshift is given by

$$\text{SFR}(M, z) = M_0 \left( \frac{M}{M_a} \right)^a \left( 1 + \frac{M}{M_b} \right)^b \left( 1 + \frac{M}{M_c} \right)^c, \quad (11)$$

where  $M_a = 10^8 M_\odot$  and  $M_0, M_b, M_c, a, b$  and  $c$  evolution with redshift are given in Table 1.

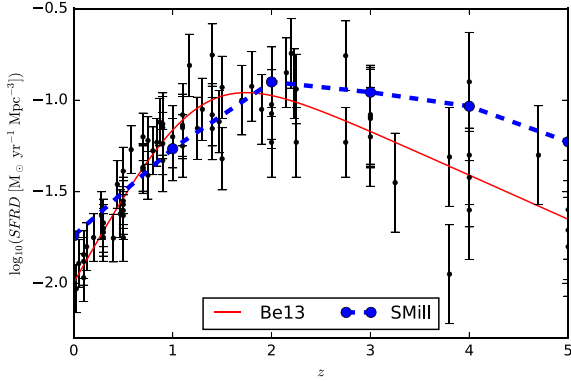
In Fig. 1, we compare the time evolution of the SFRD predicted by the two models described above. We also plot as black bold dots the data points from Behroozi et al. (2013). The split in the two models at  $z = 2$  reflects the uncertainty in this quantity due to the lack of reliable observations at higher redshifts. We thus consider Be13 as a lower estimate of the SFR while SMill will be taken as an upper estimate.

It is conventional to study the observed intensity with the quantity  $\nu I_\nu$  which we will follow in this paper. For  $\gamma = 1$ , one can use

<sup>1</sup> <http://www.peterbehroozi.com/data.html>

**Table 1.** Fit to the SFR parameters of equation (11) based on the average relations from simulated galaxy catalogues obtained by De Lucia & Blaizot (2007) and Guo et al. (2011) who post-processed the Millennium I (Springel et al. 2005) and II (Boylan-Kolchin et al. 2009) simulations, respectively.

Redshift	$M_0$	$M_b$	$M_c$	$a$	$b$	$c$
0.0	$3.0 \times 10^{-10}$	$6.0 \times 10^{10}$	$1.0 \times 10^{12}$	3.15	-1.7	-1.7
1.0	$1.7 \times 10^{-9}$	$9.0 \times 10^{10}$	$2.0 \times 10^{12}$	2.9	-1.4	-2.1
2.0	$4.0 \times 10^{-9}$	$7.0 \times 10^{10}$	$2.0 \times 10^{12}$	3.1	-2.0	-1.5
3.0	$1.1 \times 10^{-8}$	$5.0 \times 10^{10}$	$3.0 \times 10^{12}$	3.1	-2.1	-1.5
4.0	$6.6 \times 10^{-8}$	$5.0 \times 10^{10}$	$2.0 \times 10^{12}$	2.9	-2.0	-1.0
5.0	$7.0 \times 10^{-7}$	$6.0 \times 10^{10}$	$2.0 \times 10^{12}$	2.5	-1.6	-1.0



**Figure 1.** SFRD redshift evolution for the Be13 model (Behroozi et al. 2013) in solid red and for the SMill model (De Lucia & Blaizot 2007; Guo et al. 2011) in thick dashed blue. The black dots are a recollection of observational estimates of the SFRD in the literature systematized by Behroozi et al. (2013).

equations (3) and (8) and assuming that all DM haloes contribute to the signal, i.e. equation (9) to get

$$\nu \bar{I}_\nu(z) = \frac{c K(z) \text{SFRD}(z)}{4\pi(1+z)H(z)}. \quad (12)$$

We will use this quantity such that the units of  $\nu \bar{I}_\nu$  are given in  $\text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1}$ . This quantity can also be determined using LF. Defining the luminosity density as

$$\bar{\rho}_L(z) \equiv \int_{L_{\min}/L^*}^{L_{\max}/L^*} d\left(\frac{L}{L^*}\right) L \phi(L), \quad (13)$$

it follows that

$$\nu \bar{I}_\nu(z) = \frac{c \bar{\rho}_L(z)}{4\pi(1+z)H(z)}. \quad (14)$$

### 3.2 Lyman $\alpha$

The hydrogen Lyman  $\alpha$  emission line is the most energetic line coming from star-forming galaxies. Ly $\alpha$  is a UV line with a rest wavelength of 121.6 nm. It is mainly emitted during hydrogen recombinations, although it can also be emitted due to collisional excitations. As the recombination time-scale for hydrogen is usually small, we assume that the number of recombinations is the same as the number of ionizations made by UV photons emitted by young stars in star-forming galaxies. These photons have neither been absorbed by the galactic dust nor escaped the galaxy. The emission of Ly $\alpha$  photons has in fact re-processed UV emission that ionized the neutral hydrogen of the galaxy. As the Lyman  $\alpha$  photons are emitted and travel through the interstellar medium (ISM), they get absorbed

and re-emitted by neutral hydrogen until they escape the galaxy. This scattering causes the photons' direction to change randomly; hence, there is a negligible probability that photons retain their initial direction. Such a random walk in the ISM of the galaxy also increases the probability that the Ly $\alpha$  photons are absorbed by dust. The galaxy metallicity and its dust content therefore have an impact on the observed Lyman  $\alpha$  emission. Similarly, peculiar motion of the hydrogen gas broadens the line width. Thus, in an IM survey, the redshift resolution must be higher than the FWHM of the emission line at a particular redshift.

It is common in the literature to assume that Ly $\alpha$  emission is linear in the SFR (e.g. Kennicutt 1998; Ciardullo et al. 2012), i.e.  $\gamma_{\text{Ly}\alpha} = 1$ . Therefore, one can estimate the Ly $\alpha$  signal using equation (12). We model  $K^{\text{Ly}\alpha}$  in equation (8) as

$$K^{\text{Ly}\alpha}(z) = (f_{\text{dust}}^{\text{UV}} - f_{\text{esc}}^{\text{UV}}) \times f_{\text{esc}}^{\text{Ly}\alpha}(z) \times R^{\text{Ly}\alpha}, \quad (15)$$

where  $f_{\text{dust}}^{\text{UV}}$  is the fraction of UV photons that are not absorbed by dust,  $f_{\text{esc}}^{\text{UV}}$  is the fraction of UV photons that escape the star-forming galaxy,  $f_{\text{esc}}^{\text{Ly}\alpha}$  is the fraction of Ly $\alpha$  photons that escape the galaxy and  $R^{\text{Ly}\alpha}$  is a constant in units of luminosity which calibrates the line emission.

The probability of a Lyman alpha photon being emitted during a recombination is high, and assuming an optically thick ISM, case B recombination, and a Salpeter (1955) universal initial mass function we have (Ciardullo et al. 2012)

$$R_{\text{rec}}^{\text{Ly}\alpha} = 1.1 \times 10^{42} \text{ erg s}^{-1}. \quad (16)$$

Although this is the conventional relation to connect a galaxy SFR with its Ly $\alpha$  luminosity, in the low-redshift universe the use of a case A recombination coefficient might be more appropriate if the neutral gas is restricted to very dense regions. In that case, the recombination rate would be higher but the Lyman  $\alpha$  luminosity would be lower. Case A refers to the case when recombinations take place in a medium that is optically thin at all photon frequencies. On the other hand, case B recombinations occur in a UV opaque medium, i.e. optically thick to Lyman series and ionizing photons, where direct recombinations to the ground state are not allowed. For a review, please refer to Dijkstra (2014). Collisional excitations give an extra contribution that is an order of magnitude lower,  $R_{\text{exc}}^{\text{Ly}\alpha} = 4.0 \times 10^{41} \text{ erg s}^{-1}$ , at a gas temperature of  $10^4 \text{ K}$ . For now, we will ignore this contribution since it was not considered in the estimation of an observationally determined  $f_{\text{esc}}^{\text{Ly}\alpha}$ .

#### 3.2.1 UV dust absorption and escape fraction

In equation (15), the terms in parentheses correct the intrinsic luminosity for the fact that only a fraction of the UV photons produced by young stars are consumed in the ionization of ISM gas clouds. That is, a fraction of these photons escape the galaxy or are absorbed by dust. The values of  $f_{\text{esc}}^{\text{UV}}$  and  $f_{\text{dust}}^{\text{UV}}$  for a single galaxy depend on physics at both large (SFRs and gas masses) and very small scales (gas clumping). Accurate modelling of these quantities is thus out of the reach of simulations. However, observational measurements of  $f_{\text{esc}}^{\text{UV}}$  have been carried out for a few galaxies and only along a few lines of sight (Shapley et al. 2006; Siana et al. 2010; Nestor et al. 2011). Also, its estimation is dependent on measurements of the astrophysical conditions in the IGM, which on its own depends on the intensity of the UV radiation that successfully escaped the galaxy. It is therefore an indirect probe of  $f_{\text{esc}}^{\text{UV}}$  which has so far provided contradictory values. Hence, in this paper, we will consider  $f_{\text{esc}}^{\text{UV}}$  as a cosmological average of the percentage of UV photons that



escape a galaxy and  $f_{\text{dust}}^{\text{UV}}$  as a cosmological average of the percentage of UV photons that are not absorbed by dust in the galaxies. In this communication, we will assume that  $f_{\text{esc}}^{\text{UV}} = 0.2$  (Yajima et al. 2014). We will follow the conventional assumption that dust gives an average extinction of 1 mag, i.e.

$$f_{\text{dust}}^{\text{UV}} = 10^{-E_{\text{UV}}/2.5}. \quad (17)$$

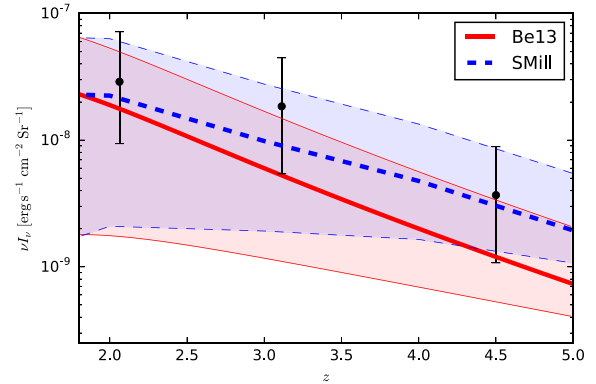
Following Sobral et al. (2013), we will consider that  $0.8 < E_{\text{UV}} < 1.2$  for our lower and higher estimates of the signal. As a first approximation, we will consider both quantities to be constant with redshift. We note that these factors fluctuate according to the galaxy properties; however, for IM studies, we are averaging over several galaxies and so the intrinsic dispersion will not be a problem.

### 3.2.2 Ly $\alpha$ escape fraction

The cosmological average of the percentage of Ly $\alpha$  photons that escape the galaxies is also very challenging to determine observationally or through simulations. The observationally based model for the escape fraction of Lyman alpha photons from Hayes et al. (2011) is given by

$$f_{\text{esc}}^{\text{Ly}\alpha} = C(1+z)^{\xi} \quad (18)$$

with  $C = (1.67^{+0.53}_{-0.24}) \times 10^{-3}$  and  $\xi = 2.57^{+0.19}_{-0.12}$ . The parameter values were estimated using observational data of the Ly $\alpha$  and H $\alpha$  relative intensities at low redshift, as well as the Ly $\alpha$  intensity relative to the continuum UV emission at high redshifts. The low values of  $f_{\text{esc}}^{\text{Ly}\alpha}$  at low redshift are likely to be due to the Ly $\alpha$  and H $\alpha$  ratio being inversely proportional to galaxy luminosity. We also note that the calibration at low redshifts was made accounting for dust absorption and possible absorption/scatter in the IGM regardless of the production mechanism of Ly $\alpha$  photons. At high redshifts, the calibration was made comparing the expected Ly $\alpha$  emission from recombinations with the observed Ly $\alpha$  emission, which included collisional excitations as a source of Ly $\alpha$ . All of these effects result in a lower estimate of  $f_{\text{esc}}^{\text{Ly}\alpha}$ . Other authors have studied the Lyman  $\alpha$  escape fraction (Dijkstra & Jeon-Daniel 2013; Ciardullo et al. 2014) arriving at similar results using LF. These are especially dependent on the LF low-end cut-off and fainter galaxies might have higher escape fractions. These estimations not only determine the escape fraction due to extinction, but also from scattering of photons in the galaxy. This is of particular importance for IM, since all emission is integrated out in a single pixel including the scattered photons that end up leaving the galaxy anyway. A significant fraction of the overall Ly $\alpha$  emission is likely to be produced by sources too faint to be observed with current technology, therefore raising the real value of  $f_{\text{esc}}^{\text{Ly}\alpha}$ . Another reason to read these results as underestimations of the escape fraction relies on the fact that such estimates are based on observations of Lyman  $\alpha$  emitters (LAEs). These are mainly sensitive to the bulge of the galaxy where the dust component is higher, and therefore have overall lower Ly $\alpha$  escape fractions. More recently, Wardlow et al. (2014) estimated a substantially higher  $f_{\text{esc}}^{\text{Ly}\alpha}$  at redshifts  $z = 2.8, 3.1$  and  $4.5$ . IM experiments have enough sensitivity to capture the Ly $\alpha$  photons scattered around the galaxy, so the Lyman  $\alpha$  escape fraction considered should only be due to dust absorption. Hence, it should be similar to the escape fraction of H $\alpha$  (see Section 3.3). This is around 40 per cent for the massive star-forming galaxies. For the less massive systems, which are out of the reach of the sensitivity of most observational experiments, it may be even higher (Dawson et al. 2012; Price et al. 2014). In light of these considerations, we will take the calibration



**Figure 2.** Average Ly $\alpha$  intensity as a function of redshift. The thick solid red line corresponds to estimates using the Be13 SFR model (Behroozi et al. 2013) and the thick dashed blue line to the estimates obtained with SMill SFR model (De Lucia & Blaizot 2007; Guo et al. 2011). The shaded regions encompass the uncertainties in  $f_{\text{dust}}^{\text{UV}}$  and  $f_{\text{esc}}^{\text{Ly}\alpha}$  and are bounded by thinner lines. The black dots were computed using the LAE LF given by Guaita et al. (2010) at  $z = 2.063$ , Ciardullo et al. (2012) at  $z = 3.113$  and Zheng et al. (2013) at  $z = 4.5$ .

of Hayes et al. (2011) as a lower estimate for the escape fraction and  $f_{\text{esc}}^{\text{Ly}\alpha} = 0.4$  constant in redshift as an upper estimate. At each redshift, we will take  $f_{\text{esc}}^{\text{Ly}\alpha} = 0.2$  as our average value.

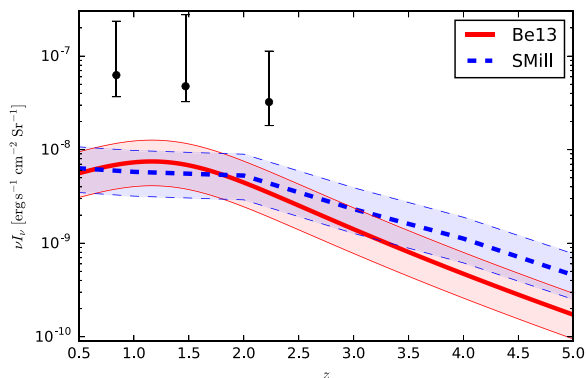
### 3.2.3 Ly $\alpha$ intensity

In Fig. 2, we compare the estimates of the average Ly $\alpha$  intensity. The lines show estimates of the expected intensity for the SFR models Be13 in red and SMill in blue computed using equation (12). We also considered variation due to the escape fraction of Lyman  $\alpha$  as well as the UV escape fraction. The dots are estimations of the average intensity computed using equation (4) and the LF (equation 5) of LAEs calibrated by Guaita et al. (2010) at  $z = 2.063$ , Ciardullo et al. (2012) at  $z = 3.113$  and Zheng et al. (2013) at  $z = 4.5$ . One should note that no further correction needs to be introduced in this estimation since LAE LF are already observational. One can see that they agree within the uncertainties considered. None the less, the LF estimates are systematically near the higher bounds of our estimates using the SFR. One of the reasons for this comes from the fact that Ly $\alpha$  escape fractions are calibrated using the emission from the bulge of galaxies which has much more extinction than the edges of the galaxy.

Ly $\alpha$  emission which arrives in the optical will be contaminated by lower redshift foregrounds as O II [373.7 nm], the O III doublet [495.9 and 500.7 nm] and other fainter metal lines as well as Balmer series lines. We will discuss the most relevant of these lines in the following subsections. Another contaminant to be taken into account is the UV continuum background emission which will be originated by young stars and quasars at higher redshifts. Although there is around one order of magnitude uncertainty, its intensity will definitely be higher than Ly $\alpha$  emission (Dominguez et al. 2011). We will not address this problem in this paper but we note that its power spectrum should be nearly flat.

### 3.3 H $\alpha$

The UV light emitted by young stars ionizes the surrounding gas in the ISM. As hydrogen recombines in a cascading process, several lines other than Ly $\alpha$  are emitted, such as H $\alpha$ . This is the lowest line



**Figure 3.** Average  $H\alpha$  intensity as a function of redshift. The thick solid red line corresponds to estimates using the Be13 SFR model (Behroozi et al. 2013) and the thick dashed blue to the estimates obtained with SMill SFR model (De Lucia & Blaizot 2007; Guo et al. 2011). The shaded regions encompass the uncertainties in  $f_{\text{esc}}^{\text{UV}}$  and  $E_{H\alpha}$  and are bounded by thinner lines. The black dots were computed using the  $H\alpha$  LF given by Sobral et al. (2013) at  $z = 0.8, 1.47$  and  $2.23$ , and corrected with  $f_{\text{esc}}^{H\alpha}$ .

of the Balmer series, with a rest-frame wavelength of 656.281 nm. Just like  $\text{Ly}\alpha$ , we will assume that the luminosity of  $H\alpha$  is linear in the SFR ( $\gamma_{H\alpha} = 1$ ), and that  $K^{H\alpha}(z)$  follows equation (15). We consider the same values and uncertainties for  $f_{\text{esc}}^{\text{UV}}$  and  $f_{\text{dust}}^{\text{UV}}$ . Kennicutt (1998) assumed an optically thick ISM, case B recombination, a Salpeter universal initial mass function and that all UV continuum is absorbed by the gas in the galaxy, arriving at

$$R^{H\alpha} = 1.3 \times 10^{41} \text{ erg s}^{-1}. \quad (19)$$

To estimate the observed  $H\alpha$  luminosity, we also need to account for dust absorption. Kennicutt (1998) assumed an extinction  $E_{H\alpha} \sim 1$  mag. Several authors have studied  $H\alpha$  extinction in galaxies and produced similar results (James et al. 2005; Sobral et al. 2013). Most  $H\alpha$  studies cannot probe the redshift evolution of  $E_{H\alpha}$ . Here we take a conservative approach and assume the extinction to be roughly constant. One should bear in mind that at higher redshifts, we expect a higher signal than predicted here, given the lower dust content. As for UV light, the  $H\alpha$  escape fraction will be given by equation (17) with  $E_{H\alpha} = 1$  mag (roughly 40 per cent) with an ‘uncertainty’ of 0.2 mag.

We compare our estimates for the average intensity of  $H\alpha$  as a function of redshift in Fig. 3. The thick solid red line shows our estimates computed using the Be13 model, while the thick blue dashed lines uses the SMill model in equation (12). The dots are estimates of the average intensity computed using equation (4) and  $H\alpha$  LF determined by Sobral et al. (2013), and corrected for extinction. One should note that these are intrinsic  $H\alpha$  LF rather than observed luminosities, as was the case for LAEs. Although one can argue that the estimates still lie within the same order of magnitude, we notice that the LF estimates are systematically higher. We are either underestimating the SFR from galaxies or the LF are overestimating the signal. Another possibility is related to the value of extinction that  $H\alpha$  photons experience. We used calibrations that are generically obtained by looking at the bulge of galaxies. These have higher extinctions than the ones experienced at the edges of the galaxy. Another possible explanation comes from the amount of UV that sources  $H\alpha$  emission which we may be underestimating. We will not explore such discrepancies further since our goal is to study the feasibility of IM with galactic emission lines.

For completeness, one should indicate which are the strongest contaminants of  $H\alpha$ . Background contamination comes mainly from oxygen lines (the  $\text{O II}$  [373.7 nm] and the  $\text{O III}$  doublet [495.9 and 500.7 nm]),  $\text{N II}$  [655.0 nm] and other lines from the Balmer series, but is mainly due to UV lines in the Lyman series. Foregrounds come mainly from continuous IR emission from galaxies.

### 3.4 $H\beta$

$H\beta$  is the second strongest line of the Balmer series, with a rest emission wavelength of 486.1 nm. In the same way as  $H\alpha$ ,  $H\beta$  follows equation (8) with  $\gamma_{H\beta} = 1$  and

$$R^{H\beta} = 4.45 \times 10^{40} \text{ erg s}^{-1}. \quad (20)$$

This is a well-known result since  $[H\beta/H\alpha] = 0.35$  for optically thick ISM, case B recombination and a Salpeter universal initial mass function. We use the same UV escape fractions as before, but assume that the  $H\beta$  extinction is slightly higher than for  $H\alpha$ , i.e.  $E_{H\beta} = 1.38$  mag (Khostovan et al. 2015). The signal estimations for  $H\beta$  are similar to the ones shown in Fig. 3 except they are rescaled by an  $[H\beta/H\alpha] f_{\text{esc}}^{H\beta} / f_{\text{esc}}^{H\alpha}$  factor.

### 3.5 Oxygen lines

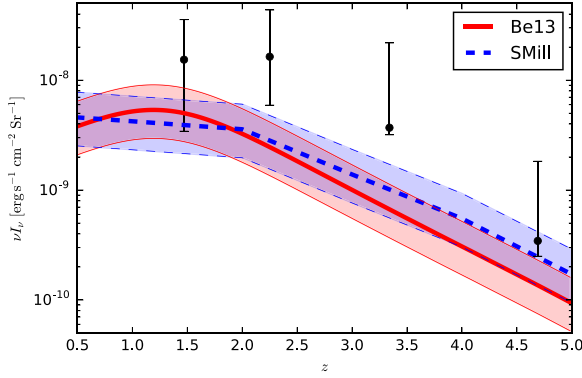
#### 3.5.1 Rest-frame emission in the optical

Ionized oxygen produces several emission lines which contribute to a galaxy spectrum in the visible and FIR. The forbidden optical oxygen line  $\text{O II}$  [372.7 nm] is a strong emission line which has been used by the Sloan Sky Digital Survey (<http://www.sdss.org>) to study galaxies and determine their redshifts. As for previous UV and optical lines, we take  $\text{O II}$  luminosity to be linearly dependent on the SFR, i.e.  $\gamma_{\text{O II}} = 1$ . The line strength has been estimated by Kennicutt (1998), assuming a Salpeter (1955) universal initial mass function and case B recombinations, as

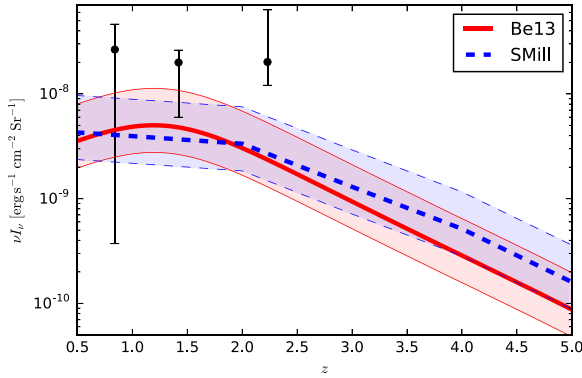
$$R^{\text{O II}} = 7.1 \times 10^{40} \text{ erg s}^{-1}. \quad (21)$$

The escape fraction of  $\text{O II}$ ,  $f_{\text{esc}}^{\text{O II}}$ , is given by equation (17) with  $E_{\text{O II}} = 0.62$  (Khostovan et al. 2015), with an extinction uncertainty of 0.2 mag. The values of  $f_{\text{esc}}^{\text{UV}}$  and  $f_{\text{dust}}^{\text{UV}}$  are the same as the ones used in previous emission lines. To compute the intensity, we used equation (3) with the mass integration range  $[10^{11}, 10^{15}] \text{ M}_{\odot}$ . We expect lower mass haloes to be metal poor (Henry et al. 2013), hence the assumed mass cut-off. In Fig. 4, we compare the different estimates of the  $\text{O II}$  signal. The thick lines show the estimates computed using Be13 (solid red) and SMill (dashed blue) models using the quoted values for  $f_{\text{esc}}^{\text{UV}}$ ,  $f_{\text{dust}}^{\text{UV}}$ ,  $f_{\text{esc}}^{\text{O II}}$  and  $R^{\text{O II}}$ . The shaded region corresponds to variations in the escape fractions, while the black dots are estimates using Khostovan et al. (2015)  $\text{O II}$  LF, and corrected for extinction. Our estimates are within the same order of magnitude as the ones using LF. This agreement is strongly dependent on the chosen minimum integration mass. Here we do not include the metallicity dependence on redshift and halo mass, since it is poorly known. This may explain the discrepancies, but our approximation still describes the trend of the signal. Such metallicity dependence is beyond the scope of this paper but one should bear in mind that this problem needs to be addressed for a proper use of  $\text{O II}$  for IM.

Other optical oxygen lines worth mentioning are the  $\text{O III}$  doublet at [500.7 nm] and [495.9 nm] with the ratio  $\text{O III}$  [500.7 nm]/[495.9 nm]  $\sim 3$ . These two lines are very hard to distinguish, and we will therefore consider the bundle of the two as



**Figure 4.** Average O II intensity as a function of redshift. The thick solid red line corresponds to estimates using the Be13 SFR model (Behroozi et al. 2013) and the thick dashed blue to the estimates obtained with SMill SFR model (De Lucia & Blaizot 2007; Guo et al. 2011). The shaded regions encompass the uncertainties in  $f_{\text{esc}}^{\text{UV}}$  and  $E_{\text{O II}}$ . The black dots were computed using the O II LF given by Khostovan et al. (2015) at  $z = 1.47, 2.25, 3.34, 4.69$ , and corrected with  $f_{\text{esc}}^{\text{O II}}$ .



**Figure 5.** Average O III intensity as a function of redshift. The solid lines correspond to the estimates of the intensity using the Be SFR model (Behroozi et al. 2013) in red and SMill SFR model (De Lucia & Blaizot 2007; Guo et al. 2011) in blue. The shaded regions encompass the uncertainties in  $R^{\text{O III}}$ ,  $f_{\text{esc}}^{\text{UV}}$  and  $E_{\text{O III}}$  and are bounded by thinner lines. The black dots were computed using the O III LF given by Khostovan et al. (2015) at  $z = 0.84, 1.42, 2.23$ , and corrected with  $f_{\text{esc}}^{\text{O III}}$ .

our ‘estimator’. We take the luminosity to be linear in the SFR and assume that  $K^{\text{O III}}$  follows equation (15) with

$$R^{\text{O III}} = (1.3_{-0.4}^{+1.2}) \times 10^{41} \text{ erg s}^{-1}. \quad (22)$$

The total luminosity of the O III lines was estimated by Ly et al. (2007) from 197 galaxies in the redshift range 0.07–1.47 in the Subaru Deep Field. The UV escape fractions are the same as before and at these wavelength, and we expect an extinction of  $E^{\text{O III}} = 1.35$  mag. We then use equation (3) to estimate the average intensity with the same mass cut-off as for O II. In Fig. 5, the thick lines correspond to our theoretical estimate of the average O III intensity, while the shaded area encompasses the variations of the O III extinction, UV escape fractions and of  $R^{\text{O III}}$ . The O III LF was determined from the O III+H $\beta$  LF calibrations of Khostovan et al. (2015) corrected by the extinction at these wavelengths and from which we subtracted the observed H $\beta$  LF. In the same way as for H $\beta$ , studying this line is important because it is a background/foreground to be cleaned from the signal, as well as being a potential line to cross-correlate with. We can see in Fig. 5 that the estimates using the SFR and LF disagree. This may be caused

either by the fact that we are considering a lower mass cut-off using the SFR or by the fact that we are overestimating the H $\beta$  contribution for the O III+H $\beta$  LF. Another plausible explanation concerns the metallicity of the galaxies which maybe lower than the one assumed to estimate  $R^{\text{O III}}$ .

### 3.5.2 Rest-frame emission in the IR

The other set of relevant oxygen lines are in the FIR. Dust clouds around star-forming regions act as re-processing bolometers and are the source of emission of O I [145  $\mu\text{m}$ ], O III [88  $\mu\text{m}$ ], O I [63  $\mu\text{m}$ ] and O III [52  $\mu\text{m}$ ]. Spinoglio et al. (2012) give a recipe to relate their luminosities to FIR luminosity. These were calibrated using several line luminosities of observed galaxies for which the IR continuum luminosity was available. Then, to relate the FIR luminosity to the SFR, we use Kennicutt (1998)

$$L_{\text{FIR}} = 2.22 \times 10^{43} \frac{\text{SFR}}{\text{M}_{\odot} \text{ yr}^{-1}} \text{ erg s}^{-1}. \quad (23)$$

In terms of equation (8), we have (Spinoglio et al. 2012)

$$\begin{aligned} K_{\text{O I}[145 \mu\text{m}]} &= 10^{39.54 \pm 0.31}, & \gamma_{\text{O I}[145 \mu\text{m}]} &= 0.89 \pm 0.06, \\ K_{\text{O III}[88 \mu\text{m}]} &= 10^{40.44 \pm 0.53}, & \gamma_{\text{O III}[88 \mu\text{m}]} &= 0.98 \pm 0.10, \\ K_{\text{O I}[63 \mu\text{m}]} &= 10^{40.60 \pm 0.17}, & \gamma_{\text{O I}[63 \mu\text{m}]} &= 0.98 \pm 0.03, \\ K_{\text{O III}[52 \mu\text{m}]} &= 10^{40.52 \pm 0.54}, & \gamma_{\text{O III}[52 \mu\text{m}]} &= 0.88 \pm 0.10. \end{aligned} \quad (24)$$

Note that these are observational fits and therefore already include extinction of the lines. Later, in Fig. 10, we show their intensities based on the previous relations and Be13 SFR model. Note that we used the same halo mass interval as for the other oxygen lines. We compared our estimates with ones using the FIR LF of Bethermin et al. (2011) and found that they are in good agreement. From Fig. 10, one can clearly see that some FIR oxygen lines are subdominant with respect to others while O I [63  $\mu\text{m}$ ] and O III [52  $\mu\text{m}$ ] are of the same order of magnitude. Since they are not good candidates to be used as prime IM tracers, we do not show a comparison with LF estimates. Instead, we only show comparisons with other FIR lines later in the paper. They none the less follow the DM distribution which, in principle, can be recovered using cross-correlations. One should also note that these FIR lines can be easily confused with each other as well as with N III [58  $\mu\text{m}$ ] and C II [158  $\mu\text{m}$ ], as is noticeable in Fig. 10.

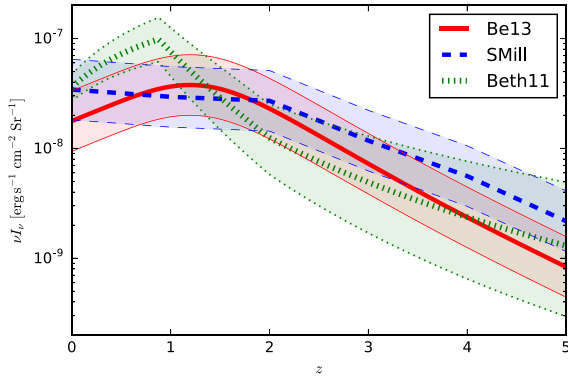
### 3.6 Ionized nitrogen

The FIR ionized nitrogen, N II [122  $\mu\text{m}$ ] and N III [58  $\mu\text{m}$ ], are cooling lines whose luminosities are obtained in the same way as for the FIR oxygen lines. From Spinoglio et al. (2012), we have

$$\begin{aligned} K_{\text{N II}[122 \mu\text{m}]} &= 10^{39.83 \pm 0.20}, & \gamma_{\text{N II}[122 \mu\text{m}]} &= 1.01 \pm 0.04, \\ K_{\text{N III}[58 \mu\text{m}]} &= 10^{40.25 \pm 0.55}, & \gamma_{\text{N III}[58 \mu\text{m}]} &= 0.78 \pm 0.10. \end{aligned} \quad (25)$$

In the same way as for the FIR oxygen lines, these fits are observational and therefore do not need to be corrected for extinction. The intensity estimates are shown in Fig. 10. As we can see, they are not suitable for IM but the considerations made for FIR oxygen lines also hold for nitrogen. We still have to understand the emission of these lines not only to have further means of probing the underlying density field (using cross-correlations), but also to clean them from the signal of other lines such as C II.





**Figure 6.** Average C II intensity as a function of redshift using equations (3) and (26). The thick solid red line corresponds to estimates using the Be13 SFR model (Behroozi et al. 2013) and the thick dashed blue to the estimates obtained with SMill SFR model (De Lucia & Blaizot 2007; Guo et al. 2011). The thick dotted green line is computed using equation (4) and the LF given by Bethermin et al. (2011). The shaded regions correspond to uncertainties in the parameters and are bounded by thinner lines.

### 3.7 C II

In most star-forming galaxies, the C II [158  $\mu$ m] transition provides the most efficient cooling mechanism for the gas in photodissociating regions. This line is also emitted from ionized regions, cold atomic gas and CO dark clouds. The C II line is therefore often the strongest IR emission line in galaxy spectra. This line has been proposed as a probe of the EoR (Gong et al. 2012; Silva et al. 2015) as well as a cosmological probe for  $0.5 < z < 1.5$  (Uzgil et al. 2014). C II emission is powered by stellar UV emission and so it correlates well with the galaxy SFR.

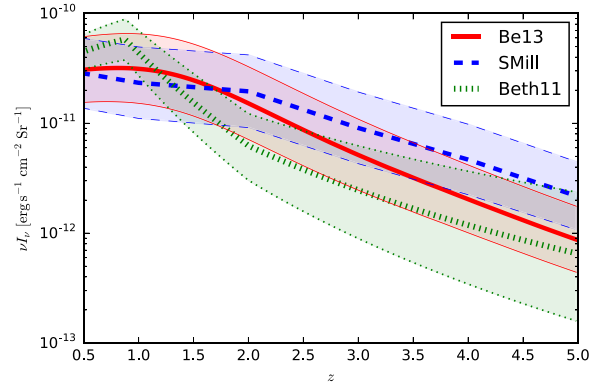
De Looze et al. (2011) made an observational fit to galaxy spectra at redshifts  $z > 0.5$  and determined the average C II luminosity to scale with the SFR as

$$L_{\text{C II}} = 5.06 \times 10^{40} \left( \frac{\text{SFR}}{\text{M}_{\odot} \text{ yr}^{-1}} \right)^{1.02} \text{ erg s}^{-1}. \quad (26)$$

From equation (8), we have  $K_{\text{C II}} = 5.06 \times 10^{40}$  and  $\gamma_{\text{C II}} = 1.02$ . This fit has a  $1\sigma$  log uncertainty of 0.27 dex. It was based on C II line luminosity measurements and on an SFR estimated from both far-UV and 24  $\mu$ m data.

The fit is, however, not appropriate for massive galaxies where the far-UV radiation field is very strong and so the preferential cooling channel is O I [63]. Also, very low mass systems are expected to be very metal poor so they should have smaller C II emission rates than the fit predicts. Still, we use equation (26) as our luminosity estimate of C II but cut off the mass integral below  $M = 10^{10} \text{ M}_{\odot}$ . In fact, the additional contribution from lower mass galaxies accounts for an increase in the intensity of at most 0.3 per cent. The mass cut-off at  $10^{10} \text{ M}_{\odot}$  was chosen because the observationally based relations used to convert between SFR and  $L_{\text{C II}}$  are consistent with emission from galaxies above  $\sim 6 \times 10^{10} \text{ M}_{\odot}$ . One should also stress that low-mass galaxies are expected to be metal poor.

In Fig. 6, we estimate the average C II intensity as a function of redshift from equations (3) and (26) using Be13 (solid red) and SMill (dashed blue) SFR models. The thick green dotted line is an estimate of the average C II intensity using the LF given by Bethermin et al. (2011) in equation (4). The shaded regions correspond to uncertainties in the estimates. We can see that our estimates are in good agreement with the estimations using the LF.



**Figure 7.** Average CO(1–0) intensity as a function of redshift using equations (3), (23), (28) and (27). The thick solid red line corresponds to estimates using the Be13 SFR model (Behroozi et al. 2013) and the thick dashed blue to the estimates obtained with SMill SFR model (De Lucia & Blaizot 2007; Guo et al. 2011). The thick dotted green line is computed using equation (4) and the LF given by Bethermin et al. (2011). The shaded regions correspond to uncertainties in the calibrations of the used expressions and are bounded by thinner lines.

Since C II is an IR line, intensity maps of the C II line will be contaminated by emission from other IR lines, namely IR oxygen and nitrogen lines as well as lines from CO rotational transitions at higher redshifts. Figs 10 and 11 make this point clearer. Although we did not study the Cl(1–0) and Cl(2–1) lines here, they are the next subdominant lines to be considered.

### 3.8 CO

Carbon monoxide (CO) rotational transitions are a powerful probe of the molecular gas in galaxies and of the astrophysical conditions in this medium. This is because the relative intensity of different transitions constrains the gas electron density and temperature. The excitation state of high CO transitions is correlated with the intensity of the radiation field exciting the molecules.

The lowest CO rotational transitions can be approximated by

$$\log_{10} (L'_{\text{CO}} [\text{K km s}^{-1} \text{ pc}^2]) = \alpha \log_{10} (L_{\text{FIR}} [\text{L}_{\odot}]) + \beta. \quad (27)$$

For normal star-forming galaxies, the CO(1–0) transition has  $\alpha_{\text{CO}(1-0)} = 0.81 \pm 0.03$  and  $\beta_{\text{CO}(1-0)} = 0.54 \pm 0.02$  (Sargent et al. 2014). The CO luminosity in  $\text{erg s}^{-1}$  can be obtained with the conversion (Carilli & Walter 2013)

$$L_{\text{CO}} = 1.88 \times 10^{29} \left( \frac{\nu_{\text{CO, rest}}}{115.27 \text{ GHz}} \right)^3 \frac{L'_{\text{CO}}}{\text{K km s}^{-1} \text{ pc}^2} \text{ erg s}^{-1}. \quad (28)$$

Note that  $\nu_{\text{CO}(1-0), \text{rest}} = 115.27 \text{ GHz}$ . In Fig. 7, we present our estimates of the CO(1–0) intensity as a function of the redshift for the two SFR models considered. In addition, we present the estimate for the intensity using the Bethermin et al. (2011) IR LF. The shaded areas indicate the uncertainties in the models, showing that the different estimates broadly agree with each other.

For CO(3–2) and CO(2–1), we use the CO ratios ( $R_{i1} \equiv L'_{\text{CO}(i-i-1)}/L'_{\text{CO}(1-0)}$ ) from Daddi et al. (2015), given by  $R21 = 0.76 \pm 0.09$  and  $R31 = 0.42 \pm 0.07$ . These held similar results to the ones shown in Fig. 7 and are consistent with the estimates using LF. Later, in Fig. 11, one can see how the intensities of the different CO lines compare.

For CO transitions (4–3) and higher, we assume that their luminosities follow the recent fit to the *Herschel* SPIRE FTS

observations of 167 local normal and starburst galaxies presented in Liu et al. (2015), i.e.

$$\log_{10}(L'_{\text{CO}}(J)[\text{K km s}^{-1} \text{ pc}^2]) = \log_{10}(L_{\text{FIR}}[L_{\odot}]) - A(J), \quad (29)$$

where  $A(J = 4) = 1.96 \pm 0.07$ ,  $A(J = 5) = 2.27 \pm 0.07$ ,  $A(J = 6) = 2.56 \pm 0.08$  and  $A(J = 7) = 2.86 \pm 0.07$ . We plot all CO rotation lines up to  $J = 7$  in Fig. 11.

### 3.9 Other lines

As a matter of completeness, it is worth leaving a note about fainter lines which will be foreground and background contaminants. In the optical range, these include other lines of the Balmer series like  $\text{H}\gamma$  at 434.1 nm, and stronger metal lines as  $\text{N II}$  [655.0 nm],  $\text{N I}$  [658.5 nm],  $\text{S II}$  [671.8 nm] and  $\text{S II}$  [673.3 nm]. The sodium and sulphur lines are expected to be stronger than the Balmer series ones. Uzgil et al. (2014) looked at fainter lines in the FIR as  $\text{Si II}$  [35  $\mu\text{m}$ ],  $\text{S III}$  [33  $\mu\text{m}$ ],  $\text{S III}$  [19  $\mu\text{m}$ ],  $\text{Ne II}$  [13  $\mu\text{m}$ ] and  $\text{Ne III}$  [16  $\mu\text{m}$ ] in conjugation with FIR oxygen lines and ionized nitrogen lines. These lines are generally weaker, but future studies will have to take them into account to properly determine the clean IM signal of stronger emission lines.

## 4 COMPARISON BETWEEN LINES

### 4.1 Measuring line intensity fluctuations

Following Visbal & Loeb (2010), the spatial fluctuations in the line signal linearly trace the DM density contrast:

$$\Delta I_v(\theta_1, \theta_2, z) \equiv I_v(\theta_1, \theta_2, z) - \bar{I}_v(z) = \bar{I}_v(z) \Delta(\mathbf{x}, z), \quad (30)$$

where  $I_v$  is the point-dependent signal and  $\bar{I}_v$  is the average line signal in a redshift bin. In the simplest scenario, neglecting redshift-space distortions, lensing and relativistic corrections,  $\Delta(\mathbf{x}, z) = \bar{b}(z) \delta(\mathbf{x}, z)$ , where  $\bar{b}$  is the luminosity-weighted bias of the emission line and  $\delta(\mathbf{x})$  is the underlying DM density contrast. The 3D power spectrum of the emission line is simply given by

$$P_I(z, k) = \bar{I}_v^2(z) \bar{b}^2(z) P_{\text{CDM}}(z, k), \quad (31)$$

where  $P_{\text{CDM}}$  is the dark matter power spectrum. Lastly, we should also estimate the shot noise contribution to each line, which is usually negligible. It is given by (Gong et al. 2011b)

$$P_{\text{shot}}(z) = \int_{M_{\min}}^{M_{\max}} dM \frac{dn}{dM} \left( \frac{L(M, z)}{4\pi D_L^2} \bar{y} D_A^2 \right)^2. \quad (32)$$

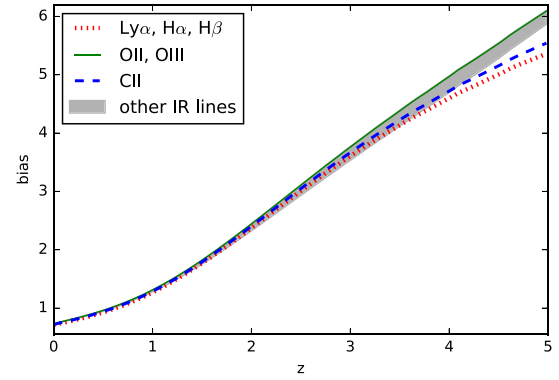
Note that contrary to threshold surveys, this shot noise is independent of the survey/instrument specifications. The luminosity-weighted bias is given by

$$\bar{b}(z) \equiv \frac{\int_{M_{\min}}^{M_{\max}} dM b(M, z) L(M, z) \frac{dn}{dM}}{\int_{M_{\min}}^{M_{\max}} dM L(M, z) \frac{dn}{dM}}, \quad (33)$$

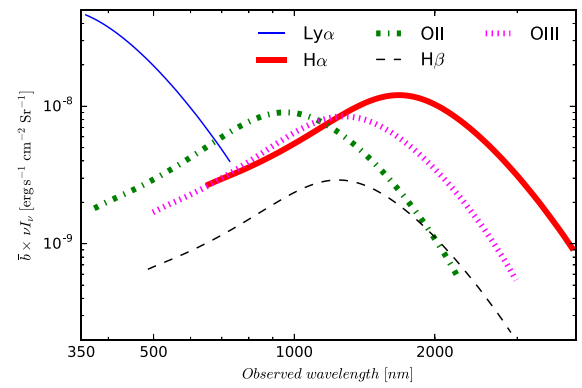
where  $b(M, z)$  is the halo bias,  $L(M, z)$  is the line luminosity and  $dN/dM$  is the halo mass function. Using this definition and taking equation (8), one can write  $\bar{b}$  in terms of the SFR:

$$\bar{b}(z, \gamma) = \frac{\int_{M_{\min}}^{M_{\max}} dM b(M, z) (\text{SFR}(M, z))^\gamma \frac{dn}{dM}}{\int_{M_{\min}}^{M_{\max}} dM (\text{SFR}(M, z))^\gamma \frac{dn}{dM}}. \quad (34)$$

Hence, the bias of each emission line depends on the mass cut-off and the value of  $\gamma$ . Without any other contribution, or dependence on the mass, one expects  $\text{Ly}\alpha$ ,  $\text{H}\alpha$  and  $\text{H}\beta$  to have the same bias.



**Figure 8.** Bias of the different lines as a function of redshift. Note that most IR lines have similar biases; hence, we show them as a shaded region. It is also clear that as we increase the mass cut-off, the bias of the lines increases.

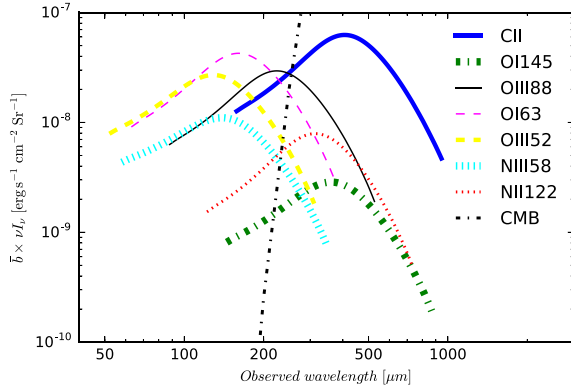


**Figure 9.** UV/optical/NIR: estimates of the product  $b \times \nu I_v$  of  $\text{Ly}\alpha$ ,  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{O II}$  and the  $\text{O III}$  doublet as a function of the observed wavelength. Due to Earth's observational constraints and the difficulties in UV observations, we only consider observed emission from the near-UV till NIR. All plotted lines are for the redshift interval  $z = 0-5$ , except  $\text{Ly}\alpha$  where we cut below  $z \sim 1.9$ .

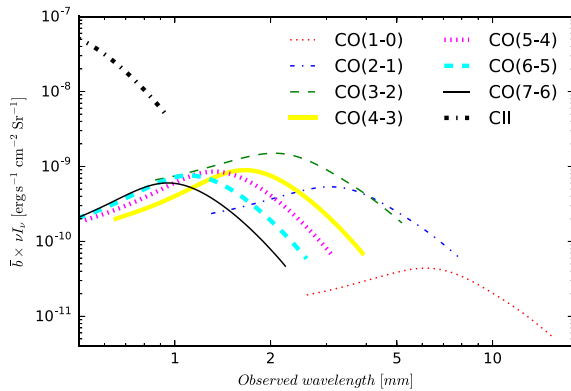
The bias difference between the hydrogen lines and the optical  $\text{O II}$  and  $\text{O III}$  doublet comes from the lower halo mass cut-off. On the other hand, the bias differences between  $\text{C II}$ , the lowest CO lines ( $J < 3$ ), the IR ionized oxygen and nitrogen lines come from the fact that they have a non-linear dependence on the SFR (i.e.  $\gamma \neq 1$ ). One can see this behaviour in Fig. 8. Note that most FIR lines have similar biases which we only represent as a shaded area. We should note that since the CO rotation lines for  $J \geq 4$  are linear in the SFR and, to first approximation, the full range of DM haloes emit these lines, we expect the bias of  $\text{CO}(J \geq 4)$  to be similar to the hydrogen emission lines. All these biases we computed using Be13 SFR model in equation (34).

### 4.2 Range of lines' dominance and contaminants

In Fig. 9, we plot the product of the bias with the estimated average intensity of the UV and optical lines studied previously.  $\text{Ly}\alpha$ ,  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{O II}$  and the  $\text{O III}$  doublet are plotted as a function of the observed wavelength up to redshift 5. Note that each line was estimated using the Be13 model. As expected, the emission from  $\text{Ly}\alpha$  is expected to be highest; this line is therefore a very strong candidate to use for IM techniques. One can also see that  $\text{O II}$  dominates for a narrow wavelength range, while  $\text{H}\alpha$  clearly dominates in the NIR. For  $z > 5$ ,  $\text{Ly}\alpha$  will be a background contaminant of  $\text{O II}$ , while for



**Figure 10.** FIR: estimates of the product  $b \times \nu I_\nu$  of C II and the other oxygen and nitrogen FIR emission lines as a function of the observed wavelength. We also plot the CMB intensity for comparison. All plotted lines are for the redshift interval  $z = 0-5$ .



**Figure 11.** COs: estimates of the product  $b \times \nu I_\nu$  CO rotational transitions as a function of the observed wavelength. All plotted CO lines are for the redshift interval  $z = 0-5$ . We also plot the C II emission for comparison.

$z \lesssim 0.2$ , H $\alpha$  will be a foreground. Similarly, O II will be a foreground for Ly $\alpha$  and will give background contamination to H $\alpha$ . As stated before, the H $\beta$  and the O III doublet will mainly be contaminants or lines to cross-correlate with. One should bear in mind that although some lines are the dominant contributor in a particular wavelength regime, the power spectrum of subdominant lines may shoot these up locally. We will not address this here and leave it for future work since such issues are highly dependent on the chosen line.

From Fig. 10, we can clearly see that C II is the dominant FIR emission line and has the potential to be used for IM after cleaning the CMB signal from the map. Still, at lower redshifts, it becomes subdominant and no other line becomes clearly dominant. In contrast, the oxygen and nitrogen lines are not the most promising lines for IM. They easily contaminate the signal of each other and C II at lower redshifts. Additionally, Cl, Si and Ne lines, which have not been included, will further contaminate the signal (Uzgil et al. 2014).

In Fig. 11, we compare the bias times the intensity of several rotational transitions of CO. We also plot the high-redshift emission of C II as comparison. We do not show the CMB emission since it will be just a background at a fixed temperature. One can see that the lowest three lines could be used for IM at intermediate to high redshifts. In the case of CO(1-0), contamination will mainly come from CO(2-1) at redshifts below  $\sim 6.2$ . Both CO(2-1) and CO(3-2)

**Table 2.** Observational wavelength range and corresponding redshift range (up to  $z = 5$ ) for which lines dominate the signal based on our estimates using the Be13 model (Behroozi et al. 2013) for the SFR. These values are to be read as indication where these lines dominate, although a wide range can still be used using masking and cleaning methods.

Emission line	Wavelength range	Redshift range	Spectral class
Ly $\alpha$	122–679 nm	0.0–4.58	UV/optical
O II	679–1131 nm	0.82–2.03	Optical/NIR
H $\alpha$	1.25–3.94 $\mu$ m	0.9–5	NIR
C II	249–948 $\mu$ m	0.57–5	FIR
CO(3–2)	1.24–4.06 mm	0.43–3.69	Radio(millimetre)
CO(2–1)	4.06–7.8 mm	2.12–5	Radio(millimetre)
CO(1–0)	8.48–15.6 mm	2.60–5	Radio(illimetre)

will be contaminated by higher  $J$  rotation lines at lower redshift and therefore cannot be used for IM in these regimes. Also, one can see that CO(3–2) will be subdominant with respect to CO(2–1) at  $z \gtrsim 3.5$  but at intermediate redshifts one can still use it as an IM probe.

Figs 9–11 merely provide an indication about which lines dominate the observed intensity and where in the spectrum. Other effects, such the power spectrum of very low redshift lines, may still create high contamination (although one expects to be able to mask it). None the less, we summarize in Table 2 the potential IM lines and the wavelength range where their intensity is dominant up to redshift 5. We also state which redshifts could be probed by these lines. These fall in a wide range of the electromagnetic spectrum, and some lie outside the observable atmospheric windows meaning space experiments are required. In summary, one can say that the best candidates to be used for cosmological IM surveys at lower to intermediate redshifts are Ly $\alpha$ , O II, H $\alpha$ , C II and the lower CO transitions.

## 5 SURVEYS

An IM experiment requires one to be able to separate the incoming light from a field into wavelength (and therefore redshift) bins. For ground-based telescopes, one can either use low-resolution spectroscopy with narrow band filters, Fabry–Pérot filters or integral field units (IFUs) for high-resolution spectroscopy. NIR lines can only be observed from space. The FIR and sub/millimetre offer more possibilities with dish experiments and optical-like settings at high altitudes. We will discuss the possibility of measuring the power spectrum with these lines for current and proposed experiments as well as feasible setups.

### 5.1 Error estimation

In this paper, we will only focus on the detectability of the 3D power spectra of different emission lines taking into account instrumental and shot noise. In this section, we will neglect the uncertainties due to the parameters used to estimate the intensity, and uncertainties due to contamination from other lines, foregrounds and backgrounds.

For an experiment with sensitivity  $\sigma_N$ , the noise power spectrum is given by

$$P_N = \sigma_N^2 \times V_{\text{pixel}}, \quad (35)$$

where  $V_{\text{pixel}}$  is the comoving volume corresponding to the redshift and angular resolution of the considered experiment. The simplest estimate of the error in measuring  $P(k)$  is

$$\Delta P(k_j) \simeq \frac{P_T(k_j)}{\sqrt{N_k(k_j)}}, \quad (36)$$

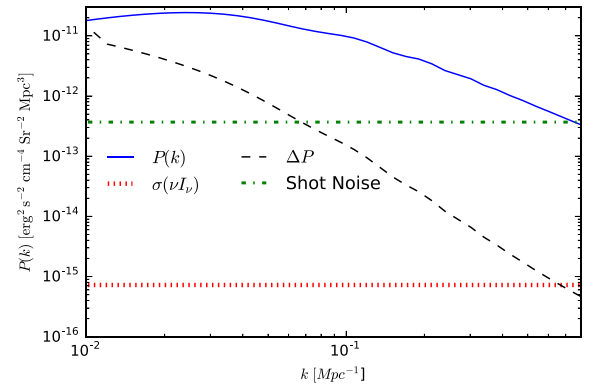
where  $P_T = P_S + P_N + P_{\text{shot}}$  is the total power in a scale  $k_j$ , while  $N_k$  is the number of accessible modes at a scale  $k$ . For a 3D survey  $N_k(k_j) = k_j^3 \Delta k V_{\text{sample}} / 2\pi^2$ , where  $V_{\text{sample}}$  is the comoving volume of the survey and  $\Delta k$  is the chosen  $k$ -bins. This approach is only valid within a regime  $k \in [k_{\min}, k_{\max}]$ , where  $k_{\min} \sim 2\pi/L$  is given by the smallest side of the sample volume and  $k_{\max} \sim 2\pi/\Delta L$  by the biggest side of the resolution pixel. Outside this range, the  $k$ -space is approximately 2D, i.e.  $N_k(k_j < k_{\min}) = k_j \Delta k S_{\text{sample}} / 2\pi$ . For a single scale (bin  $j$ ), the signal-to-noise is given by  $\text{SNR}(k_j) = P_S(k_j) / \Delta P(k_j)$ . Adding up all the signal-to-noise, for a given experiment one has

$$\text{SNR}^2 = \sum_j \left( \frac{P_S(k_j)}{\Delta P(k_j)} \right)^2. \quad (37)$$

## 5.2 Ly $\alpha$ IM

An experiment using IFUs is the HETDEX – Hobby–Eberly Telescope Dark Energy eXperiment (Hill et al. 2008, [www.hetdex.org](http://www.hetdex.org)), a 3 year survey designed to see approximately 0.8 million LAEs in the redshift range  $z \sim 1.9$ – $3.5$  covering  $300 \text{ deg}^2$  with a filling factor of  $1/4.5$ . This will be achieved using the instrument VIRUS (Visible Integral-field Replicable Unit Spectrograph) which is composed of 150 wide-field IFUs, each with 224 optical fibres. Pairs of IFUs are built as single units so it will observe 33 600 pixels, obtaining a spectrum for each pixel. HET has an FoV of 22 arcmin in diameter while VIRUS will only provide a coverage of around  $1/4.5$  of the full FoV. Each fibre has a diameter of 1.5 arcsec but there will be a dithering pattern that will effectively give an angular resolution of  $\delta\Omega_{\text{pixel}} \simeq 9.05 \text{ arcsec}^2 \simeq 2.13 \times 10^{-10} \text{ sr}$ . A single field will be observed for 1200 s using three separated dithering exposures, giving a total survey time of 1200 h (assuming 140 observation nights over 3 years). VIRUS will have a wavelength coverage from 3500 to 5500 Å with a spectral resolution of  $\lambda/\Delta\lambda = 800$  or an average 6.4 Å wavelength resolution. The quoted line sensitivity for 20 min of integration time at  $z = 2.1$  is  $1.28 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$  or  $\sigma(\nu I_\nu) = 6.02 \times 10^{-8} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$ .

An IM experiment with VIRUS will not need to use the full resolution of the experiment since we can consider larger IM pixels and wavelength bins. By doing so, one can reduce the experimental noise and increase the detectability of the signal. As an example, the given resolution of 6.4 Å in the range 3500–5500 Å gives around 300 redshift bins of size  $\Delta z \sim 0.005$ . Such redshift resolution is unnecessary for cosmological studies as well as potentially introducing errors due to the emission line profile of Lyman  $\alpha$ . For IM, the width of the bin needs to be bigger than the observed FWHM of Ly $\alpha$  line. Yamada et al. (2012) find a rest-frame Ly $\alpha$  FWHM which is smaller than 1 Å. Even taking a conservative approach, at redshift 3 one only expects the observed FWHM to be  $\sim 4$  Å. Similarly, small angular scales are not good probes of the cosmology since they are polluted by uncertainties in the clustering of matter in DM haloes. Although changing the pixel size could increase the sensitivity in the pixel, this does not alter the noise power spectrum (equation 35) since the decrease in the sensitivity is cancelled by the increase in the pixel volume. The only variable that one could



**Figure 12.** Estimated power for Ly $\alpha$  IM (solid blue), HETDEX noise (thick dotted red), Ly $\alpha$  shot noise (dash-dotted green) and the error in estimating the power spectrum (dashed black) at  $z = 2.1$ .

change in HETDEX is the integration time. We can then rewrite the instrumental noise for IM as

$$\sigma_N^{\text{HETDEX}} = \sigma_{\nu I_\nu} \sqrt{\frac{20 \text{ min}}{\delta t}}, \quad (38)$$

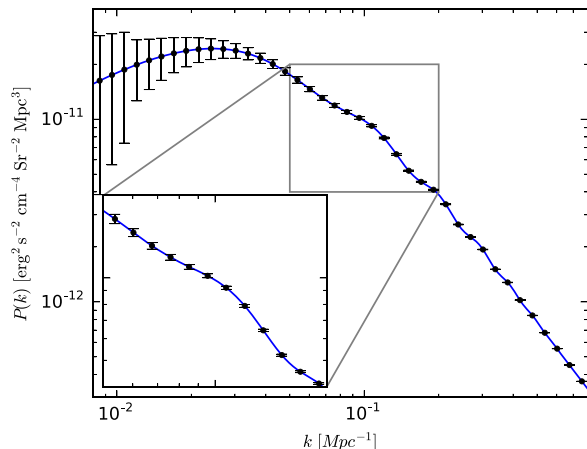
where  $\delta t$  is the new integration time per pointing. One should note that for a 20 min integration, the sensitivity is up to an order of magnitude higher than the expected signal (compare with Fig. 2).

Note that  $\delta t$  is just the ratio between the total observation time  $t_{\text{TOT}}$  and the number of pointings  $N_p$ , where  $N_p$  is the ratio between the surveyed area  $S_{\text{Area}}$  and FoV<sub>VIRUS</sub>. If one takes into account the 1 min of overheads, the integration time is  $\delta t = t_{\text{TOT}}[\text{min}] / N_p - 1 \text{ min}$ . The HETDEX quoted numbers point to an integration time of 20 min for roughly 3500 pointings. Unfortunately, this is for a filling factor of 4.5. If we keep the sparse sampling of the HETDEX survey, then both the instrumental and shot noise need to increase by the same amount as the filling factor (Chiang et al. 2013).

In Fig. 12, we show the power spectrum of Ly $\alpha$  IM (solid blue), the power coming from HETDEX instrumental noise (dotted red), Ly $\alpha$  shot noise (dot-dashed green) and  $\Delta P_{\text{Ly}\alpha}$  (dashed black) at  $z = 2.1$ . We assume a sparse coverage of  $300 \text{ deg}^2$  and a sample with  $\Delta z = 0.4$  around  $z = 2.1$ . One can see that HETDEX is so sensitive that the instrumental noise is subdominant with respect to Ly $\alpha$  shot noise (although this is dependent on assumptions), while both are orders of magnitude lower than the power spectrum *per se*. This becomes clearer when plotting the power spectrum with the estimated error bars, as in Fig. 13. Not only would one measure the Ly $\alpha$  power spectrum very well but one would also have good enough statistics to resolve the wiggles from the baryon acoustic oscillations (BAOs). Ly $\alpha$  IM is therefore a good candidate to study further including all models for background continuum emission, foregrounds and contaminants. We will leave this to a future paper.

HETDEX is primarily a galaxy survey, but can be used as a Ly $\alpha$  intensity mapper as we have pointed out. We have also shown that Ly $\alpha$  intensity shot noise is much higher than the HETDEX instrumental noise, if it is used as an IM experiment. One should note that Ly $\alpha$  intensity shot noise is intrinsic to the line (although it is affected by the sampling coverage). Hence, as an IM survey, HETDEX does not gain anything from longer integration times. In fact, this causes the survey to be sparse, thus increasing the shot noise, and reducing the area covered, i.e. increasing the error bars on the power spectrum due to a lower number of available  $k$ -modes.





**Figure 13.** Ly $\alpha$  IM power spectrum at  $z = 2.1$  with forecasted error bar for HETDEX.

Therefore, one concludes that if HETDEX was an IM experiment, one ought to consider larger survey areas with less time per pointing in order to maximize the SNR.

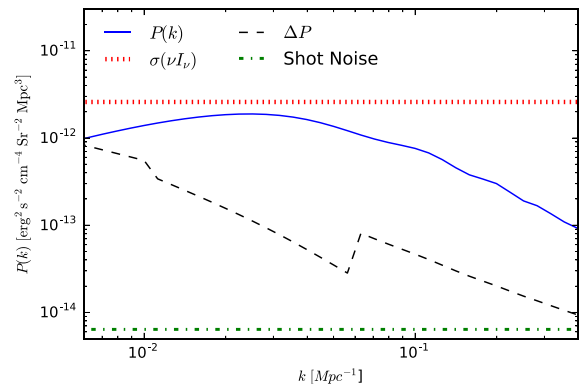
### 5.3 H $\alpha$ IM

H $\alpha$  is an optical line at rest but will be observed in the IR for  $z \gtrsim 0.06$ . The Earth's observational window in the IR is reduced to a minor collection of narrow observational gaps in the NIR. Hence, we need to make space observations, such as *Euclid* (Amendola et al. 2013, <http://www.euclid-ec.org>) which will use H $\alpha$  emission for its galaxy survey.

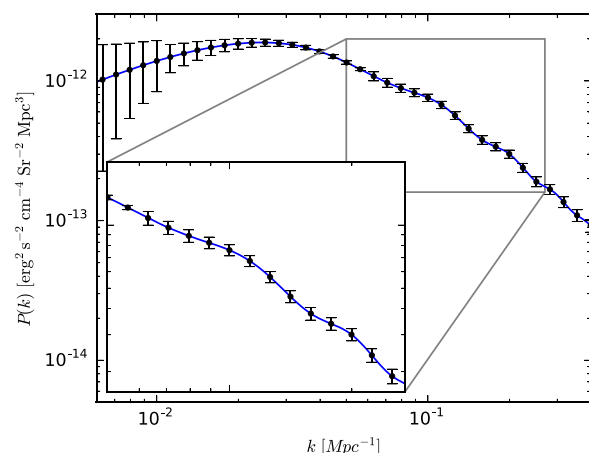
Here we will focus on the planned space telescope *SPHEREx* (Doré et al. 2014, <http://spherex.caltech.edu>) and revisit their IM mode. *SPHEREx* is an all-sky space telescope in the NIR having four linear variable filters. Although it covers the full sky, only  $\sim 7000 \text{ deg}^2$  will be of any cosmological use. Its spectral resolution is  $\lambda/\Delta\lambda = 41.5$  for  $0.75 < \lambda < 4.1 \mu\text{m}$  and  $\lambda/\Delta\lambda = 150$  for  $4.1 < \lambda < 4.8 \mu\text{m}$ . The instrument has a pixel size of  $6.2 \text{ arcsec} \times 6.2 \text{ arcsec} = 9.03 \times 10^{-10} \text{ sr}$ . The  $1\sigma$  flux sensitivity depends on the wavelength bin in consideration but remains within the same order of magnitude. For H $\alpha$  at  $z = 1.9$ , we will take  $\sigma(\nu I_\nu) \sim 1 \times 10^{-6} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$ .

*SPHEREx* is a space survey so we cannot change the integration time to improve the sensitivity of the experiment. The size of the pixel is only relevant for the maximum  $k$  that is accessible. In fact, changing the pixel size will not alter the noise power as we saw before. As a thought experiment, let us start by analysing how well we would observe the power spectrum at  $z = 1.9$  with a sample size of  $\Delta z = 0.4$ . In Fig. 14, we show as the blue solid line the H $\alpha$  power spectrum at  $z = 1.9$ , as dotted red the instrumental noise power spectrum, as dot-dashed green the H $\alpha$  shot noise and as the black dashed line the error in measuring the power spectrum. The number of modes mainly comes from the 2D information encoded in the surveyed area. We can see that we should be able to measure the power spectrum on large scales with *SPHEREx*, as well as the BAOs. This becomes clearer in Fig. 15.

A more futuristic experiment is the Cosmic Dawn Intensity Mapper (Cooray et al. 2016). It would cover a part of the spectrum slightly broader than *SPHEREx* ( $\lambda = 0.7\text{--}7 \mu\text{m}$ ) and would have a flux sensitivity 30–50 times higher than *SPHEREx*. This would have a clear impact on the error bars of the power spectrum, especially at larger scales. Although Cosmic Dawn Intensity Mapper



**Figure 14.** Estimated power for H $\alpha$  IM (solid blue), *SPHEREx* noise (thick dotted red), H $\alpha$  shot noise (dash-dotted green) and the error in estimating the power spectrum (dashed black) at  $z = 1.9$ .

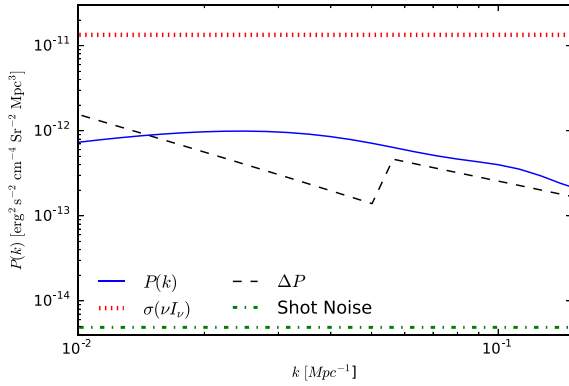


**Figure 15.** H $\alpha$  IM power spectrum at  $z = 1.9$  with forecasted error bar for *SPHEREx*.

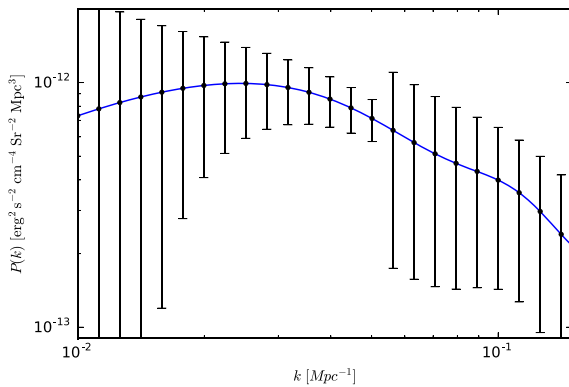
was proposed to study EoR, it can, and should, have a commensal H $\alpha$  (and O II) IM survey at intermediate to low redshifts.

### 5.4 O II IM

O II is mainly an optical and an NIR line, and from Fig. 9 we see that there is a short-wavelength window where we expect it to be dominant. This broadly corresponds to the transition between optical and NIR. Although one could still use ground telescopes with filters or a spectrograph for O II IM, we will also consider *SPHEREx* (Doré et al. 2014) for this line. We take the same instrumental settings as the ones described for H $\alpha$ , but adapt the sensitivity to the required wavelength. For O II at  $z = 1.2$ , we will take  $\sigma(\nu I_\nu) \sim 3 \times 10^{-6} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$ . Since we are looking at a different range in the spectrum, the redshift resolution shifts to  $\delta z = 0.05$ . We will assume that the sample has  $\Delta z = 0.4$ . Since we are intrinsically looking at lower redshifts, the voxel is smaller in comparison with H $\alpha$ . We therefore have less modes as one can see in Fig. 16. Similarly, instrumental noise is dominant over the power spectrum. With this setting, one finds that it will be hard to measure the O II 3D power spectrum using IM. We present the results in Fig. 17, where it is clear that due to the low volume of the voxel (in comparison with previous lines) we cannot measure the BAO wiggles. There are not enough  $k$ -modes to overcome the instrumental noise.



**Figure 16.** Estimated power for O II IM (solid blue), *SPHEREx* noise (thick dotted red), O II shot noise (dash-dotted green) and the error in estimating the power spectrum (dashed black) at  $z = 1.2$ .



**Figure 17.** O II IM power spectrum at  $z = 1.2$  with forecasted error bar from *SPHEREx*.

O II thus requires more sensitive experiments with improved angular resolution. This could, in principle, be done by optical ground telescopes equipped with IFUs. To our knowledge, no such telescope exists in this wavelength range. As pointed out previously at the end of Section 5.3, the proposed Cosmic Dawn Intensity Mapper (Cooray et al. 2016) has an improved flux sensitivity that can go from 30 to 50 times higher than *SPHEREx*. With such improvement, we naively expect the error bars in Fig. 17 to shrink by the square of the same factor.

### 5.5 C II IM

We saw in Section 4.2 that approximately above redshift 0.6, C II becomes the dominant line in the FIR and submillimetre part of the spectrum. In this regime, one requires radio dishes or telescopes equipped with bolometers at very high altitudes to overcome absorption from the atmosphere.

A radio instrument looking at these frequencies is ALMA (<https://almascience.eso.org>). ALMA is located in Chajnantor plateau at 5000 m above sea level and has been constructed to give insights into the birth of stars and stellar systems. It is primarily an interferometer; however, in order to cover the required area, one would need to make observations in single-dish mode. The lowest redshifts are not accessible since only bands 7 and 8 allow observation in single-dish mode. Hence, we focus on C II emission from  $z = 2.8$  to 5. One can find the details for these bands in Table 3.

**Table 3.** Experimental details of ALMA.

Band	Wavelength range (mm)	FoV diameter (arcsec)	$T_{\text{sys}}$ (K)
1	6.7–9.6	139.7	26
3	2.6–3.6	54.2	60
7	0.8–1.1	16.9	219
8	0.6–0.8	12.6	292

The temperature rms in a single pixel is given by

$$\sigma_T = \frac{T_{\text{sys}}}{\sqrt{2N_d \Delta\nu \Delta t}}, \quad (39)$$

where  $T_{\text{sys}}$  is the system temperature,  $N_d$  is the number of used dishes,  $\Delta\nu$  is the frequency resolution and  $\Delta t$  is the integration time. In this regime, we can use the Rayleigh–Jeans approximation to estimate the intensity rms as

$$\sigma_{I_\nu} = \frac{2\nu^2 k_B}{c^2} \sigma_T. \quad (40)$$

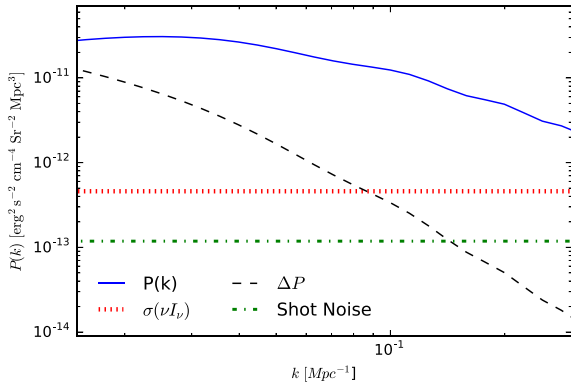
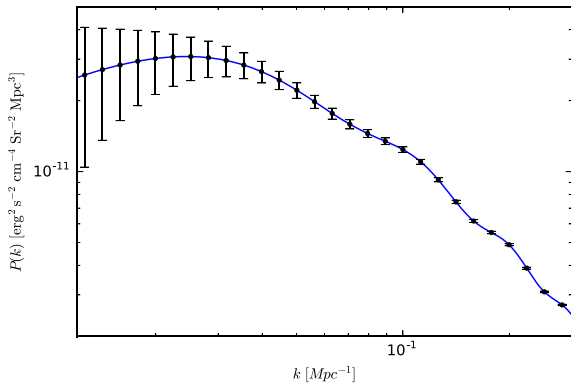
ALMA has a frequency resolution that can go from 3.8 kHz to 25 MHz which can be tuned accordingly. We compute the 3D power spectrum at  $z = 3$  using band 8 and  $z = 4.5$  using band 7. For both, we consider  $\Delta z = 0.4$  and  $\delta z = 0.05$ . Note that we would have to bundle several low-resolution bins together for this redshift binning. Due to ALMA constraints, only four antennas can work in single-dish mode together. We assume 1 h of integration per pointing. Unfortunately, ALMA cannot do these kinds of studies since the errors are several orders of magnitude higher than the power spectrum itself. One would need to lower the system temperature and add more antennas working in single-dish mode.

Another possibility is to consider a C II IM experiment inspired by TIME (Staniszewski et al. 2014), which is a proposed C II IM for EoR. Since we want to study the late universe with IM experiments, we need to adapt the dish size and angular resolution for the targeted line. In Table 4, we summarize the assumed experimental details. For full details of TIME and TIME-*Pilot*, please refer to Staniszewski et al. (2014) and to Crites et al. (2014). Note that we assume that the number of bolometers is the same. We also considered a higher noise since we expect the noise to increase with frequency. For this experiment, we decided to target emission from a central redshift of  $z = 2.2$  and assumed a sample with  $\Delta z = 0.4$  and a survey of 100 deg<sup>2</sup>. This is already enough for the instrumental noise power spectrum to go well below the expected power spectrum, as we can see in Fig. 18. This would be as good as HETDEX for Ly $\alpha$  but looking at a complementary tracer of the DM distribution. The low noise level becomes clear when we plot the 3D power spectrum with forecasted error bars in Fig. 19. One should note that these IM experiments have low angular resolution, but it is still enough to see the BAO wiggles in the power spectrum. We do not show a zoomed-in part of the power spectrum because the TIME-like C II IM experiment is merely a proposal, in opposition to HETDEX (ongoing) and *SPHEREx* (to be approved). Still we want to stress that a C II experiment with this configuration could measure the BAOs and give a complementary probe of the large-scale distribution of matter in the universe.

We have not exhausted the range of possible experiments for C II IM. One should still refer to the work of Uzgil et al. (2014), which considers using an atmospheric balloon or a cryogenic satellite to perform C II IM. Satellite experiments have the clear advantage of not suffering from FIR absorption from moisture. Still, ground

**Table 4.** TIME-like experimental details.

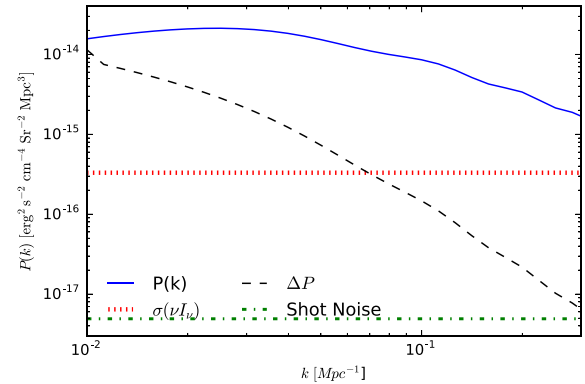
Line	Dish size (m)	Band (GHz)	$\delta\Omega$ (arcmin <sup>2</sup> )	Survey area (deg <sup>2</sup> )	NEFD (erg s <sup>-1</sup> cm <sup>-2</sup> Hz <sup>-1</sup> $\sqrt{s}$ )	FoV (arcmin <sup>2</sup> )	Total time (h)	$\delta\nu$ (GHz)
C II	6	500–700	0.4 <sup>2</sup>	100	$20 \times 10^{-26}$	$0.4 \times 25.6$	2000	0.4
CO(3–2)	10	100–200	1 <sup>2</sup>	250	$5 \times 10^{-26}$	$1 \times 64$	2000	0.4

**Figure 18.** Estimated power for C II IM (solid blue), TIME-like experiment noise (thick dotted red), C II shot noise (dash-dotted green) and the error in estimating the power spectrum (dashed black) at  $z = 2.2$ .**Figure 19.** C II IM power spectrum at  $z = 2.2$  with forecasted error bar for a TIME-like experiment.

experiments at high altitudes can largely minimize such issues, as ALMA does.

## 5.6 CO

We saw in Section 4.2 that the three lowest CO rotational lines are the best candidates to be used as IM tracers of the underlying DM distribution. Currently, COPPS (Keating et al. 2016) is looking for the CO(1–0) cosmological signal to recover the power spectrum. Similarly, COMAP (Li et al. 2016) will be looking at a window from 30 to 34 GHz, hence looking at CO(1–0) emission from  $z = 2.4$ – $2.8$  and CO(2–1) emission at late EoR redshifts,  $z = 5.8$ – $6.7$ . Although it will have high spectral resolution, it will cover a very limited area of the sky. Since these experiments already exist with the specific goal of doing IM of CO(1–0), we will focus on other CO rotation lines. Another reason not to look at this line is that synchrotron emission from galaxies and jets from quasars peak at these frequencies. Therefore, let us focus, for now, on the other two CO rotational

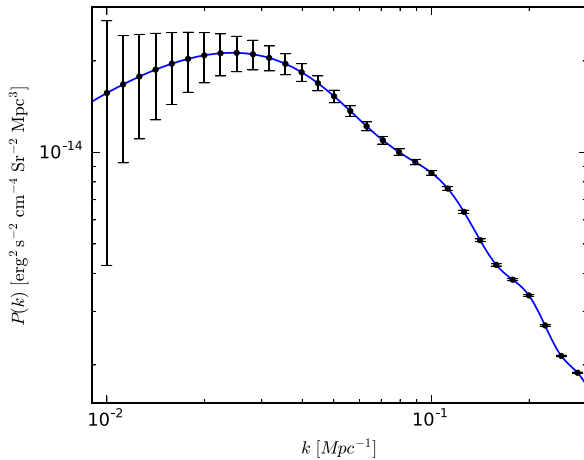
**Figure 20.** Estimated power for CO(3–2) IM (solid blue), TIME-like experiment noise (thick dotted red), CO(3–2) shot noise (dash-dotted green) and the error in estimating the power spectrum (dashed black) at  $z = 2.0$ .

lines, CO(3–2) and CO(2–1). Bands 4 and 1 of ALMA (see Table 3) will look at these frequencies, respectively. Unfortunately, ALMA performs as badly as for CO as it does for C II IM.

The last experimental setup that we consider is again a TIME-like experiment to perform CO(3–2) IM. We change the experimental setting to a 10 m dish which will give around 1 arcmin angular resolution, as one can see in Table 4. This lower resolution allows for a larger survey area for the same survey time, which we take to be 250 deg<sup>2</sup>. As a test example, we consider a sample survey with  $\Delta z = 0.4$  around a central redshift of  $z = 2.0$  and with a redshift resolution of  $\delta z = 0.01$ . Fig. 20 shows the power spectrum of CO(3–2) IM (solid blue), the power coming from the TIME-like instrumental noise (dotted red), CO(3–2) shot noise (dot-dashed green) and  $\Delta P_{\text{CO(3–2)}}$  (dashed black) at  $z = 2.0$ . Although the instrumental noise is higher than the 3D power spectrum, there are enough  $k$ -modes such that  $\Delta P$  falls below the power spectrum itself. This becomes clearer in Fig. 21 where we can see that one could even measure the BAOs.

## 6 IGM EMISSION

In the circumgalactic medium (CGM), the gas is warm/hot and at low redshifts one expects it to be considerably metal rich. This medium can be observed because of its far-UV and X-ray emission lines. However, the strength of these emission lines is considerably smaller than ISM emission in the same part of the spectrum. Also, given the large pixels involved in IM studies, the power spectrum from CGM emission follows that of galaxies and so it has little impact for the proposed large-scale studies. In the colder IGM filaments, the composition of the gas is close to primordial. Therefore, its main radiative cooling channels should be emission in the Ly $\alpha$  and the H $\alpha$  lines. The intensity of these lines from recombinations in IGM filaments should be of the order of 10 per cent of the emission from galaxies and its power spectra should be relatively flat compared to the emission from galaxies in the same lines



**Figure 21.** CO(3-2) IM power spectrum at  $z = 2.0$  with forecasted error bar for a TIME-like experiment.

(Silva et al. 2013; Silva, Kooistra & Zaroubi 2016). Given the large cross-section of photons of the Lyman series, there will be absorption of UV background photons (mainly between the Lyman  $\beta$  and Lyman  $\alpha$  frequencies) in the neutral hydrogen gas contained in the cold filaments. These photons will be re-emitted as Ly $\alpha$  photons, which follow the spatial distribution of the gas filaments. The importance of this emission depends on the average density of the hydrogen gas in the filaments, which includes gas in clouds with lower column densities than the gas probed by the Lyman alpha forest. It also depends on the intensity of the UV background in the relevant frequency range and at the relevant redshifts, which is only known within one order of magnitude at most. Consistently, modelling the intensity of the scattered radiation requires detailed simulations and the use of radiative transfer codes, which goes beyond the scope of this study. The intensity of this emission should however be smaller than the emission from galaxies. Also, the smaller bias of IGM emission and the flatness of its spectra will, in any case, make it subdominant relative to the galaxies' power spectra in the post-reionization epoch.

## 7 CONTAMINATION IN IM

This study indicates the best lines/redshifts for performing IM studies, assuming that the foregrounds of each line could be efficiently removed. Unfortunately, this is not always the case and so a prior detailed study of a line's foreground/background contamination is required. Only then, one can seriously use these lines as cosmological probes. The contamination study should be made for a specific target line and take into account the proposed experimental setup, the available foreground removal methods, the target spatial range in the power spectra and the allowed error for the proposed cosmological/astrophysical goals.

This indeed sets the way for future work in this field. To properly access the feasibility of all these cleaning techniques, one has to use simulations of the 3D distribution of haloes. For each halo in a line of sight, one attributes a luminosity from each line (including all subdominant lines) as prescribed in Section 3, as well as a continuum emission. Similarly, one should assume a continuous extragalactic background. Then, by redshifting all the emission to the observers frame, one should access how well one can disentangle the signal assuming background smoothness and cross-correlations.

This possibility not only opens new windows to probe the distribution of matter in the universe but also to study global properties of galaxies and their astrophysical properties. Similarly, different lines will be present in different halo ranges and redshift ranges. These studies open the possibility to use IM to study the chemical evolution of the universe as well as the HMF.

### 7.1 Line contamination

IM will suffer contamination from other lines in the foreground or background of the target sources, which are observed at the same frequency range. Given that in IM studies sources are not resolved, one cannot easily differentiate between the different emission lines. Contamination from interloping lines will show up in the observed power spectrum projected into the redshift of the target line. The power spectra of foreground lines will have a higher contamination with respect to what is suggested by Figs 9–11 (Gong et al. 2014). As a result, even in the cases when the intensity of the target line is dominating the signal, it will usually be necessary to use a cleaning method to deal with the bright line foregrounds. It has however been shown, in several previous studies, that for an IM experiment with a resolution similar to the proposed experiments described in Section 5, the required masking of bright sources will only cover a small percentage of the pixels in the observational maps. The success of this masking procedure is due to the contamination being highly dominated by a small number of bright foreground sources and so it can be reduced by a few orders of magnitude with this procedure. Also, the signal and the foreground positions are uncorrelated and so for small masking percentages the power of the target signal will practically not be attenuated.

None the less, an efficient method of dealing with line foregrounds/backgrounds is to cross-correlate the target line with the galaxies' number density or with other line emission at the same redshift. This can, in principle, be used to recover the power spectrum of the target line free of contamination at a first-order approximation. This technique is not only powerful to clean contaminants but also to increase the significance of the detection of overdensities of emission.

### 7.2 Continuum background and foreground contamination

Additional contamination will come from the extragalactic continuum background. This background originates from AGN emission and continuum stellar, free-free, free-bound and two-photon emission from star-forming galaxies, and it can be estimated as in Cooray et al. (2012). This radiation can be fitted out from intensity maps due to its smooth evolution in frequency compared to the line intensity fluctuations. Besides its smoothness, one can use cross-correlations between different emission lines, which suffer contamination from different parts of the spectrum, to at first order obtain a clean signal. Although this is generically true, the background contamination will always be present in the power spectrum error determination. Hence, one can immediately foresee that large areas with enough  $k$ -modes are needed to reduce cross-correlation errors.

Continuum foregrounds in intensity maps, of the UV and optical lines studied here, are not as well known as the continuum foregrounds in the radio band. These have been intensely studied in the context of 21 cm IM. However, the continuum contamination relative to the intensity of the signal from the target line should be much smaller for UV and optical lines than that for radio lines (even if it is still dominating the signal) and so the residuals from the removal of these contaminants are less problematic.



**Table 5.** Survey details in estimating  $P(k)$  for the different lines in consideration.

	Line	Area (deg <sup>2</sup> )	$z$	$\delta z$	$\delta\Omega$ (10 <sup>-9</sup> sr)	$\Delta z$	$P_N$ [erg <sup>2</sup> s <sup>-2</sup> cm <sup>-4</sup> sr <sup>-2</sup> Mpc <sup>3</sup> )	$k$ -range (Mpc <sup>-1</sup> )	SNR
HETDEX	Ly $\alpha$	300	2.1	0.005	0.213	0.4	$7.24 \times 10^{-16}$	0.009–0.3	489
SPHEREx	H $\alpha$	7000	1.9	0.07	0.903	0.4	$2.59 \times 10^{-12}$	0.009–0.3	105
	O II	7000	1.2	0.05	0.903	0.4	$1.34 \times 10^{-11}$	0.02–0.1	11
TIME-like	C II	100	2.2	0.002	13.5	0.4	$4.59 \times 10^{-13}$	0.01–0.3	294
		2 (50 h)					$3.67 \times 10^{-13}$	0.05–0.3	42
	CO(3–2)	250	2.0	0.01	85	0.4	$3.31 \times 10^{-16}$	0.01–0.3	471
		2 (50 h)					$1.06 \times 10^{-16}$	0.05–0.3	45

## 8 DISCUSSION

IM studies open new windows to probe not only the distribution of matter in the universe but also to study global properties of galaxies. Particularly IM is useful to probe the BAO scale independently from galaxy surveys. The evolution of the HMF is also an interesting cosmological goal although it will be difficult to disentangle it from the evolution of astrophysical properties of galaxies. On the astrophysical side, IM of metal lines can be used to probe the poorly known chemical evolution of the gas in the universe.

In Table 5, we summarize the experimental and sample details for each line we consider for IM. In the last two columns, we present the SNR for each survey for a fixed  $k$ -range. Although higher SNRs can be attained going to smaller scales, we caped at  $k = 0.3 \text{ Mpc}^{-1}$  so we do not need to worry about issues with non-linear scales. Also, below these scales, the voxel of IM may be too small for the assumption that astrophysical fluctuations between galaxies average out. From the last column of Table 5, we see that using HETDEX as a Ly $\alpha$  IM survey would be the best performing survey out of the ones considered in this paper. Although H $\alpha$  IM with SPHEREx is possible, its SNR is not as good as for Ly $\alpha$  or the FIR lines with a TIME-like experiment. The proposed O II survey will have the worst performance of the IM experiments considered, given the smaller physical volume to be observed. This is due to the relatively low redshifts of the observational frequency window where O II is the dominant emission line. C II with a TIME-like experiment also has a high SNR but it is still lower than CO(3–2). This is mainly due to the fact that for a fixed survey time a CO(3–2) IM survey can sample a bigger area of the sky. In fact, a CO(3–2) IM survey can do as well as Ly $\alpha$  with HETDEX. Since TIME-like experiments are in the realm of possibilities, we also performed the calculations for a pilot experiment. For both C II and CO(3–2), we computed the SNR for a smaller 50 h survey covering 2 deg<sup>2</sup> of the sky. These can be used as a proof of concept despite the lower SNR. Looking at Table 5, one can say that Ly $\alpha$  and CO(3–2) would be the best lines to perform IM. This is not necessarily true due to contamination issues. As said before, different lines are sensitive to different mass ranges of the HMF, to the chemical evolution of the Universe and to the underlying astrophysical processes within galaxies. Hence, we would like to stress that all lines should be seen as complementary, although Ly $\alpha$  is the best starting point.

This work is intended to probe the new windows opened by line IM for studying our cosmology. On top of looking at the distribution of matter in the Universe, one can study the astrophysical properties of galaxies. We started by modelling the luminosity of a line in terms of halo mass. We proposed a general prescription based on previous theoretical and observational studies. Although one expects scatter in these relations on small scales, it should average out in big enough voxels. In our models, we find Figs 9–11 which indicate which are

the best candidates for line IM. Unsurprisingly, these are Ly $\alpha$ , O II, H $\alpha$ , C II and the lowest rotation lines of CO. We then follow by estimating power spectrum error bars for each line with reasonable experimental settings for a fixed redshift range. We find that one can measure the power spectrum of these lines, assuming a cleaned signal. Unlike 21 cm emission of H I, UV, optical and IR lines have stronger line contaminants that need to be considered. We discussed foreground contamination and ways to extract the target signal from observational maps in Section 7 and found that there are promising ways to successfully clean many of these maps. The literature already has a wide discussion on such methods either for the EoR (Visbal & Loeb 2010; Lidz et al. 2011; Silva et al. 2013, 2015; Gong et al. 2014; Breyse, Kovetz & Kamionkowski 2015) or for the late universe (Croft et al. 2016; Lidz & Taylor 2016), which we will explore in the future.

This work considers the 3D power spectrum  $P(k)$ , as it is conventional in galaxy surveys. None the less, one should point out that IM allows for tomographic studies using the angular power spectra  $C_\ell$ . There are several advantages on doing so. First of all,  $P(k)$  is gauge dependent and suffers from projection effects, which does not happen when using  $C_\ell$ . Furthermore, the contributions from lensing and the so-called GR corrections to number count fluctuations (Yoo 2010; Bonvin & Durrer 2011; Challinor & Lewis 2011) and temperature/intensity fluctuations (Hall, Bonvin & Challinor 2013) are easily included in the angular power but not in the 3D  $P(k)$ .

We note that in order to successfully do science with line intensity maps, it will be necessary to perform further detailed studies. Galaxy emission studies will have to be done for foreground/interloping lines. Also, continuum contamination estimates should be made accounting for dust absorption and/or propagation in the IGM. This is needed to properly access the feasibility of all contamination cleaning techniques. As prescribed in Section 3, one should use cosmological simulations of the 3D distribution of haloes and attribute line luminosities and galactic continuum emission to each on the top of a continuous extragalactic continuum. Then, by redshifting all the emission to the observers' frame, one should assess how well one can disentangle the signal, accounting for masking rates of interloping lines, the frequency smoothness of the continuum contamination and cross-correlations. These are general guidelines since the ideal method depends on the target line.

We conclude by emphasizing that IM, with lines other than H I, shows great potential to measure the large-scale distribution of matter in the Universe. This is especially true since independent sets of lines can measure the BAO scale around  $z = 2$ , a further test of the expansion of the universe at intermediate redshifts. One should also note that these lines acquire an important role in the context of the multi-tracer technique (McDonald & Seljak 2009; Seljak 2009) to beat cosmic variance. As Fonseca et al. (2015) showed, the

improvement in measuring large-scale effects is greatly increased as the ratio of the two biases deviates from  $\sim 1$ . Hence, these lines can give a better bias ratio for a particular pairing of DM tracers.

## ACKNOWLEDGEMENTS

We would like to thank Eiichiro Komatsu for useful discussions. We thank Imogen Whittam for proof reading the paper. JF and MGS acknowledge support from South African Square Kilometre Array Project and National Research Foundation. MBS thanks the Netherlands Foundation for Scientific Research support through the VICI grant 639.043.006. MGS also thanks the support of FCT under the grant PTDC/FIS-AST/2194/2012. AC acknowledges support from the NSF grant AST-1313319 and the NASA grants NNX16AF39G and NNX16AF38G.

## REFERENCES

- Alonso D., Ferreira P. G., 2015, *Phys. Rev. D*, 92, 063525
- Amendola L. et al., 2013, *Living Rev. Relativ.*, 16, 6
- Ansari R. et al., 2012, *A&A*, 540, A129
- Bacon D. et al., 2015, *Proc. Sci.*, Synergy between the Large Synoptic Survey Telescope and the Square Kilometre Array. SISSA, Trieste, PoS (AASKA14)145
- Bagla J., Khandai N., Datta K. K., 2010, *MNRAS*, 407, 567
- Battye R. A., Davies R. D., Weller J., 2004, *MNRAS*, 355, 1339
- Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, *MNRAS*, 434, 1239
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Bethermin M., Dole H., Lagache G., Borgne D. L., Penin A., 2011, *A&A*, 529, A4
- Bigot-Sazy M.-A. et al., 2016, in Qain L., ed., *ASP Conf. Ser. Vol. 502*, Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015. Astron. Soc. Pac., San Francisco, p. 41
- Bonvin C., Durrer R., 2011, *Phys. Rev. D*, 84, 063505
- Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, *MNRAS*, 398, 1150
- Breyse P. C., Kovetz E. D., Kamionkowski M., 2014, *MNRAS*, 443, 3506
- Breyse P. C., Kovetz E. D., Kamionkowski M., 2015, *MNRAS*, 452, 3408
- Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, *ApJ*, 803, 21
- Camera S., Santos M. G., Ferreira P. G., Ferramacho L., 2013, *Phys. Rev. Lett.*, 111, 171302
- Carilli C., Walter F., 2013, *ARA&A*, 51, 105
- Challinor A., Lewis A., 2011, *Phys. Rev. D*, 84, 043516
- Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, *Phys. Rev. Lett.*, 100, 091303
- Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, *Nature*, 466, 463
- Chen X., 2012, *Int. J. Mod. Phys.: Conf. Ser.*, 12, 256
- Chiang C.-T. et al., 2013, *J. Cosmol. Astropart. Phys.*, 1312, 030
- CHIME Collaboration, 2012, CHIME Overview, Available at [http://chime.phas.ubc.ca/CHIME\\_overview.pdf](http://chime.phas.ubc.ca/CHIME_overview.pdf)
- Ciardullo R. et al., 2012, *ApJ*, 744, 110
- Ciardullo R. et al., 2014, *ApJ*, 796, 64
- Cooray A., Gong Y., Smidt J., Santos M. G., 2012, *ApJ*, 756, 92
- Cooray A. et al., 2016, preprint ([arXiv:1602.05178](https://arxiv.org/abs/1602.05178))
- Crites A. T. et al., 2014, *Proc. SPIE*, 9153, 91531W
- Croft R. A. C. et al., 2016, *MNRAS*, 457, 3541
- Daddi E. et al., 2015, *A&A*, 577, A46
- Dawson W. A. et al., 2012, *ApJ*, 747, L42
- Dawson K. S. et al., 2016, *AJ*, 151, 44
- de Jong R. S. et al., 2012, *Proc. SPIE*, 8446, 84460T
- De Looze I., Baes M., Bendo G. J., Cortese L., Fritz J., 2011, *MNRAS*, 416, 2712
- De Lucia G., Blaizot J., 2007, *MNRAS*, 375, 2
- Dijkstra M., 2014, *Publ. Astron. Soc. Aust.*, 31, 40
- Dijkstra M., Jeesson-Daniel A., 2013, *MNRAS*, 435, 3333
- Dominguez A. et al., 2011, *MNRAS*, 410, 2556
- Doré O. et al., 2014, preprint ([arXiv:1412.4872](https://arxiv.org/abs/1412.4872))
- Fernández X. et al., 2016, *ApJ*, 824, L1
- Flaugher B., 2005, *Int. J. Mod. Phys. A*, 20, 3121
- Fonseca J., Camera S., Santos M. G., Maartens R., 2015, *ApJ*, 812, L22
- Gong Y., Cooray A., Silva M. B., Santos M. G., Lubin P., 2011a, *ApJ*, 728, L46
- Gong Y., Chen X., Silva M., Cooray A., Santos M. G., 2011b, *AJ*, 740, L20
- Gong Y., Cooray A., Silva M., Santos M. G., Bock J., Bradford C. M., Zemcov M., 2012, *ApJ*, 745, 49
- Gong Y., Cooray A., Santos M. G., 2013, *ApJ*, 768, 130
- Gong Y., Silva M., Cooray A., Santos M. G., 2014, *ApJ*, 785, 72
- Guaita L. et al., 2010, *ApJ*, 714, 255
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Hall A., Bonvin C., Challinor A., 2013, *Phys. Rev. D*, 87, 064026
- Hayes M., Schaerer D., Ostlin G., Mas-Hesse J. M., Atek H., Kunth D., 2011, *ApJ*, 730, 8
- Henry A., Martin C. L., Finlator K., Dressler A., 2013, *ApJ*, 769, 148
- Hill G. J. et al., 2008, in Kodama T., Yamada T., Aoki K., eds, *ASP Conf. Ser. Vol. 399*, Panoramic Views of Galaxy Formation and Evolution. Astron. Soc. Pac., San Francisco, p. 115
- James P. A., Shane N., Knapen J., Etherton J., Percival S., 2005, *A&A*, 429, 851
- Keating G. K., Marrone D. P., Bower G. C., Leitch E., Carlstrom J. E., DeBoer D. R., 2016, *ApJ*, 830, 34
- Kennicutt R. C. J., 1998, *ARA&A*, 36, 189
- Kerp J., Winkel B., Ben Bekhti N., Flöer L., Kalberla P. M. W., 2011, *Astron. Nachr.*, 332, 637
- Khostovan A. A., Sobral D., Mobasher B., Best P. N., Smail I., Stott J. P., Hemmati S., Nayyeri H., 2015, *MNRAS*, 452, 3948
- Laureijs R. et al., 2011, *ESA-SRE*, p. 12
- Levi M. et al., 2013, preprint ([arXiv:1308.0847](https://arxiv.org/abs/1308.0847))
- Li T. Y., Wechsler R. H., Devaraj K., Church S. E., 2016, *ApJ*, 817, 169
- Lidz A., Taylor J., 2016, *ApJ*, 825, 143
- Lidz A., Furlanetto S. R., Oh S. P., Aguirre J., Chang T.-C., Dore O., Pritchard J. R., 2011, *ApJ*, 741, 70
- Liu D., Gao Y., Isaak K., Daddi E., Yang C., Lu N., van der Werf P., 2015, *ApJ*, 810, L14
- Loeb A., Wyithe J. S. B., 2008, *Phys. Rev. Lett.*, 100, 161301
- LSST Dark Energy Science Collaboration, 2012, preprint ([arXiv:1211.0310](https://arxiv.org/abs/1211.0310))
- Ly C. et al., 2007, *ApJ*, 657, 738
- McDonald P., Seljak U., 2009, *J. Cosmol. Astropart. Phys.*, 0910, 007
- Mellema G. et al., 2013, *Exp. Astron.*, 36, 235
- Nestor D. B., Shapley A. E., Steidel C. C., Siana B., 2011, *ApJ*, 736, 18
- Newburgh L. B. et al., 2016, preprint ([arXiv:1607.02059](https://arxiv.org/abs/1607.02059))
- Price S. H. et al., 2014, *ApJ*, 788, 86
- Pullen A., Chang T.-C., Dore O., Lidz A., 2013, *ApJ*, 768, 15
- Pullen A. R., Dore O., Bock J., 2014, *ApJ*, 786, 111
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Santos M. et al., 2015, *Proc. Sci.*, Cosmology from a SKA HI intensity Mapping Survey. SISSA, Trieste, PoS (AASKA14)019
- Sargent M. T. et al., 2014, *ApJ*, 793, 19
- Schlegel D., White M., Eisenstein D., 2009, in *Astronomy*, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey
- Seljak U., 2009, *Phys. Rev. Lett.*, 102, 021302
- Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Adelberger K. L., 2006, *ApJ*, 651, 688
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Siana B. et al., 2010, *ApJ*, 723, 241
- Silva M., Santos M. G., Gong Y., Cooray A., 2013, *ApJ*, 763, 132
- Silva M. B., Santos M. G., Cooray A., Gong Y., 2015, *ApJ*, 806, 209
- Silva M. B., Kooistra R., Zaroubi S., 2016, *MNRAS*, 462, 1961
- Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, *MNRAS*, 428, 1128
- Spergel D. et al., 2015, preprint ([arXiv:1503.03757](https://arxiv.org/abs/1503.03757))

- Spinoglio L., Dasyra K. M., Franceschini A., Gruppioni C., Valiante E., Isaak K., 2012, *ApJ*, 745, 171
- Springel V. et al., 2005, *Nature*, 435, 629
- Staniszewski Z. et al., 2014, *J. Low Temp. Phys.*, 176, 767
- Switzer E. R. et al., 2013, *MNRAS*, 434, L46
- Uzgil B. D., Aguirre J. E., Bradford C. M., Lidz A., 2014, *ApJ*, 793, 116
- Visbal E., Loeb A., 2010, *J. Cosmol. Astropart. Phys.*, 1011, 016
- Visbal E., Haiman Z., Terrazas B., Bryan G. L., Barkana R., 2014, *MNRAS*, 445, 107
- Wardlow J. L. et al., 2014, *ApJ*, 787, 9
- Yajima H., Li Y., Zhu Q., Abel T., Gronwall C., Ciardullo R., 2014, *MNRAS*, 440, 776
- Yamada T., Matsuda Y., Kousai K., Hayashino T., Morimoto N., Umemura M., 2012, *ApJ*, 751, 29
- Yoo J., 2010, *Phys. Rev. D*, 82, 083508
- Zheng Z.-Y. et al., 2013, *MNRAS*, 431, 3589

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.